ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

Scuola di Scienze
Dipartimento di Fisica e Astronomia
Corso di Laurea Magistrale in Fisica

Statistical Mechanics Perspectives on Boltzmann Machines

Relatore:
Prof. Pierluigi Contucci

Correlatore:
Dott. Emanuele Mingione

Presentata da:
Francesco Camilli

Anno Accademico 2018/2019

Abstract

Questa tesi contiene un'analisi fisico-matematica di sistemi disordinati, con particolare attenzione ai modelli di campo medio. Sebbene siano usualmente concepiti come una semplificazione, questi ultimi possono rivelarsi ardui da trattare in maniera formale, come testimoniano i trent'anni di sforzi compiuti per dimostrare la correttezza della teoria di Giorgio Parisi sul modello di Sherrington e Kirkpatrick (SK).

Al fine di introdurre le tecniche principali utilizzate in tutto il lavoro, il primo capitolo è stato dedicato al modello di Curie-Weiss. In esso oltre aii risultati classici [17], si presenta una nuova proprietà di stabilità rispetto a piccole perturbazioni nella normalizzazione dei termini di accoppiamento spin-spin nell'Hamiltoniana. Il modello viene risolto per interazioni ferromagnetiche ed antiferromagnetiche.

Il secondo capitolo contiene una serie di risultati sul già menzionato modello SK, i cui accoppiamenti sono estratti da una gaussiana standard $\mathcal{N}(0,1)$ ed indipendenti. Per quest'ultimo, si provano l'esistenza del limite termodinamico e la correttezza del *replica symmetry breaking ansatz* di Parisi per l'energia libera con l'ausilio di due bound. La dimostrazione del bound superiore va oltre gli scopi di questo lavoro, ed è presentata soltanto in maniera sintetica. Il bound inferiore invece è provato in dettaglio tramite lo schema di Aizenmann, Sims e Starr [2][3].

Nei due capitoli successivi, sono stati studiati modelli in cui l'invarianza per permutazioni fra gli spin è preservata soltanto all'interno di alcuni loro sottogruppi. Il nostro interesse si volge inizialmente verso i casi deterministici di cui il multi-layer, una particolare istanza dei precedenti, viene esplicitamente risolto con un metodo nuovo che utilizza bound dall'alto e dal basso. A seguire, un'analisi di modelli in cui la matrice di interazioni tra le specie è definita (negativa o positiva). Per sistemi multi-specie disordinati si mostra la soluzione per il caso ellittico, descritto da una matrice delle covarianze delle interazioni definita positiva [7][19].

Il caso iperbolico, cioé SK multi-layer, chiamato anche *Deep Boltzmann Machine* (DBM), è discusso nel capitolo 5. A causa dell'iperbolicità, in questo caso si riesce solo a fornire un bound dall'alto per l'energia libera, costruito con un'opportuna combinazione di energie libere di SK. Questo risultato si può usare per lo studio delle regioni di *annealing* e di *replica symmetry*, due regimi associati a fasi di alte temperature, in linea di principio differenti. Mostriamo che, a campo esterno nullo, la stabilità della soluzione *replica symmetric* è implicata dalle stesse condizioni che assicurano l'*annealing* [4]. Per concludere, si mostra che le dimensioni relative dei layers della DBM possono essere scelte in maniera da comprimere al meglio la regione di *annealing*.

Contents

Introduction

This thesis deals with the rigorous approaches to the statistical mechanics of disordered systems, in particular those that have emerged from Giorgio Parisi's work on the mean field theory of the spin glass phase. From the beginning of the eighties to the present times the efforts to transform that theory into a set of mathematical statements with proofs has produced spectacular results, still they have achieved only a part of such purpose.

Technically, the topic of this work is a probability measure called *quenched*, made of the classical Boltzmann prescription and successively averaged on the disorder noise. The starting point of our work is the simplest, yet very profound when treated with mathematical rigour, classical mean field model with ferromagnetic or antiferromagnetic interactions.

Chapter 1 is indeed dedicated to an introduction to the Curie-Weiss model. Beside the classical rigorous results [17] we investigate a novel property of stability under normalisation that entails useful mathematical properties to compute thermodynamic quantities. In this chapter, both the ferromagnetic and anti-ferromagnetic versions of the model are considered, and their free-energy in the thermodynamic limit is computed.

In chapter 2, the celebrated Sherrington-Kirkpatrick model [23] is studied. The latter is still a mean field model, but the couplings between the spins are *i.i.d.* sampled from a standard gaussian $\mathcal{N}(0,1)$. The existence of the thermodynamic limit is proved, thanks to Guerra-Toninelli's interpolation, together with the famous Guerra's bound for the free energy [14]. We will presented the latter as a consequence of the Aizenmann-Sims-Starr's extended variational principle [2] [3]. As widely known, the other bound was found by Michel Talagrand only in 2006 [26], confirming once and for all the long-standing conjecture due to Giorgio Parisi. The detailed proof of Talagrand's bound lays beyond the scope of the present work, nevertheless some hints will be provided [18].

The following three chapters move from the usual setting with the global permutation group symmetry to the case in which the model itself has a weaker property, and therefore a richer physical content, of invariance only under a subgroup of it.

Chapter 3 describes some multi-species deterministic models [6][12][13], namely those models in which the permutation symmetry of the spins holds only in restricted subgroups. Here, the importance of the first chapter will be clear. In fact, the free energy of every model studied herein shares many characteristics with the Curie-Weiss one. The multi-layer Curie Weiss model is also considered and solved. The interest in the latter is motivated by the fact that the interactions draw a graph equal to that of deep networks.

Chapter 4 contains one of the main topics of the present work: disordered multi-species models [7]. As done in chapter 3, we break the permutation symmetry of the spins, leaving it untouched only inside some subgroups, the species. Here we will deal with *elliptic* models, namely those in which the interactions among spins of the same species are dominating over the inter-species ones. A Replica Symmetry Breaking (RSB) *ansatz* for the free energy is provided, and proved to be correct [19]. Here a multi-species version of the Aizenmann-Sims-Starr extended variational principle is proposed.

In chapter 5, the reader will find a study of the annealed and replica symmetric regions of the SK model. These regions coincide for a vanishing external field, under the assumption that the celebrated Almeida-Thouless (AT) line is correct [27], though it has not been proved yet. The mentioned AT line is believed to separate the so called RSB region from the replica symmetric one, in which a simple expression of the free energy holds. Finally, it follows an analysis of the Deep Boltzmann Machine (DBM), based on a work in preparation [4] by A. Barra, D. Alberici, P. Contucci and E. Mingione. In Statistical Mechanics language, the DBM is a multi-layer SK model, namely a particular instance of a hyperbolic multi-species model. Because of this hyperbolicity, we are able only to build an upper bound to the free energy with an appropriate combination of SK free energies. The previous result is then used to find the phase space region, characterized by the temperature and the relative layers sizes, in which the annealed and quenched free energies of the DBM coincide.

Furthermore, we propose a possible replica symmetric approximation for the DBM. In analogy with the SK model, imposing the stability of the replica symmetric solution with vanishing external field, we find a set of conditions that coincide with those ensuring the annealed solution. Then, we show that by an appropriate tuning of the parameters of the model we can squeeze the annealed region.

Notations

$\Sigma_N = \{-1, 1\}^N$; configuration space of an N spin system;

$H_N(\sigma)$: Hamiltonian of an N spin system;

$Z_N(\beta) = \sum_{\sigma \in \Sigma_N} e^{-\beta H_N(\sigma)}$: partition function at inverse absolute temperature β;

$p_N(\beta) = \frac{P_N(\beta)}{N} = \frac{1}{N} \log Z_N(\beta)$: Generating functional of the connected moments of the hamiltonian density or pressure per particle;

$\omega_N(\cdot) = Z_N^{-1} \sum_{\sigma \in \Sigma_N} e^{-\beta H_N(\sigma)}(\cdot)$: expectation w.r.t. Boltzmann-Gibbs measure;

$\omega_0(\cdot) = \frac{1}{2^N} \sum_{\sigma \in \Sigma_N} (\cdot)$: expectation w.r.t. uniform measure on the configuration space;

$H_N(t)$, $p_N(t)$, $Z_N(t)$: interpolating hamiltonian, pressure and partition function;

$\omega_{N,t}(\cdot) = Z_N(t)^{-1} \sum_{\sigma \in \Sigma_N} e^{-\beta H_N(t)}(\cdot)$: expectation w.r.t. the measure induced by $H_N(t)$;

$\mathbb{E}_x$: expectation w.r.t. the r.v. x;

$m_N(\sigma) = \frac{1}{N} \sum_{i=1}^N \sigma_i$: configurational magnetization;

e_N : ground state energy;

$q_N(\sigma, \tau) = \frac{1}{N} \sum_{i=1}^N \sigma_i \tau_i$: overlap of two spin configurations $\sigma, \tau \in \Sigma_N$;

H_σ : gaussian family of Hamiltonians, labeled by $\sigma \in \Sigma_N$;

$\mathbb{E}$: generic expectation w.r.t. the disorder;

$W_N(\sigma; \beta, h) = e^{\beta h \sum_{i=1}^N \sigma_i}$: deterministic weight;

$\Omega^{(2)}_{N,t}(\cdot) = Z_N(t)^{-2} \sum_{\sigma,\tau\in\Sigma_N} e^{-\beta(H_\sigma(t)+H_\tau(t))}(\cdot):$ two replica Boltzmann-Gibbs expectation w.r.t. the measure induced by the interpolating Hamiltonian;

$r = (\xi_\alpha.p_{\alpha,\alpha'}):$ ROSt with ξ_α random weights and $p_{\alpha,\alpha'}$ overlap kernel;

$G_{r,M}(\beta,h):$ Cavity functional, with ROSt r;

$\bar{\xi}_\alpha:$ Ruelle probability cascades random weights;

$p_{RS}(\beta,h,\bar{q}):$ replica symmetric ansatz in its minimum point;

$\mathcal{P}(x,\beta,h):$ Parisi replica symmetry breaking ansatz;

$(\mathbf{m})_{s=1,\dots,K}$, $(\mathbf{h})_{s=1,\dots,K}:$ K species configurational magnetization, local external field;

$\Delta:$ reduced interaction matrix for multi-species deterministic models;

$\Delta^2:$ covariance matrix of the random interactions, disordered multi-species models;

$q_s(\sigma,\tau) = \frac{1}{N_s}\sum_{i\in\Lambda_s}\sigma_i\tau_i:$ s-species overlap;

$\mathcal{P}(x,\beta,\mathbf{h}):$ RSB ansatz for elliptic multi-species models;

$p^{an}, p^A:$ annealed pressure per particle;

$\stackrel{\text{iid}}{\sim}$: identically, independently distributed;

$\stackrel{\text{D}}{=}$: equality in distribution or in law.

Chapter 1

Curie-Weiss model

The Curie-Weiss model is a mean field deterministic spin model. As such, it is characterized by the fact that each spin interacts with all the others with the same coupling J. In addition one could add a magnetic field that biases the spins towards one of the two possible directions of magnetization. The CW model is one of the simplest models exhibiting a phase transition. As we shall see later, above certain values of the inverse absolute temperature β, the spins align and remain aligned even when the magnetic field is turned off.

Historically speaking, the thermodynamic limit of this model was found long ago, but a rigorous proof of the existence of it, without knowing the limit itself, was given only in the late 90's. There are several ways for this simple model to prove the existence of the thermodynamic limit and all of them rely on the sub- or super-additivity of the free energy.

The first proof makes use of a *symmetrization lemma* (see [11]) which is applicable only if the interaction term in the hamiltonian has a specific normalization, as discussed in the following. The search for a more general proof will lead us to formulate the celebrated interpolation method, due to Guerra and Toninelli [17], that turns out to be applicable even to the disordered version of the CW model, *i.e.* the Sherrington-Kirkpatrick model discussed in Chapter 2. We also show here, that a slight change of normalization in the interacting part of the hamiltonian, needed to give the desired extensive behaviour to the latter, that does not affect the thermodynamic limit, can instead affect the so called *finite size corrections* to the free energy.

For convenience, in this chapter and the ones that will follow, we deal with the generating functional of the connected moments of the Hamiltonian, instead of the free energy. This functional is commonly called *pressure* in mathematical physics jargon, not to be confused with the usual pressure. If F_N is the free energy, when N

particles are considered, the pressure is simply:

$$-\beta F_N = P_N = \log \sum_{\sigma\in\Sigma_N} e^{-\beta H_N(\sigma)} \tag{1.1}$$

1.1 Existence of the thermodynamic limit

1.1.1 Sub- and super-additivity

The existence of the limit can be proved in at least two ways. Both of them aim to prove that the pressure of the system verifies the two hypothesis of the following simple lemma.

Lemma 1.1.1 (Sub-additivity (super-additivity) or Fekete's lemma). *Let $(a_n)_{n\in\mathbb{N}\cup\{0\}}$ be a sequence of real numbers. If $\exists\, \underline{c}$ (or $\bar{c}$) $\in \mathbb{R} : n\underline{c} \leq a_n$ ($a_n \leq n\bar{c}$) $\forall n \in \mathbb{N}$ and the sequence is sub-additive (respectively super-additive) $a_{n_1+n_2} \leq a_{n_1} + a_{n_2}$ ($a_{n_1+n_2} \geq a_{n_1} + a_{n_2}$) $\forall n_1, n_2 \in \mathbb{N}$, then the limit $\lim_{n\to\infty} a_n/n$ exists and:*

$$\lim_{n\to\infty} \frac{a_n}{n} = \inf_{n\in\mathbb{N}} \frac{a_n}{n} \quad \left(= \sup_{n\in\mathbb{N}} \frac{a_n}{n}\right) \tag{1.2}$$

Proof. To begin with, let us observe that:

$$\frac{a_{2n}}{2n} \leq \frac{a_n + a_n}{2n} = \frac{a_n}{n} \tag{1.3}$$

It follows by induction that the same inequality is true for every non negative integer k as shown here:

$$\frac{a_{kn}}{kn} \leq \frac{a_{(k-1)n} + a_n}{kn} \leq \frac{ka_n}{kn} = \frac{a_n}{n} \tag{1.4}$$

If we take a pair of natural numbers $b, d \in \mathbb{N} : b > d$, we can always find another pair n, r, with $n \in \mathbb{N}$, $r \in \mathbb{N}\cup\{0\}$, such that: $b = nd + r$. By iterating the previous inequality we get:

$$\frac{a_b}{b} = \frac{a_{nd+r}}{nd+r} \leq \frac{a_{nd}}{nd+r} + \frac{a_r}{nd+r} \leq \frac{na_d}{nd+r} + \frac{a_r}{b} \tag{1.5}$$

Observe that a_n/n is bounded by the constants $\underline{c}$ and $\bar{c}$ from below and from above respectively. This implies that lim sup and lim inf are finite. Thus one can safely take the $\limsup_{n\to\infty}$ of both sides. The inequality will be preserved, and b approaches infinity too:

$$\limsup_{b\to\infty} \frac{a_b}{b} \leq \frac{a_d}{d} \quad \forall d \in \mathbb{N} \tag{1.6}$$

After taking the $\inf_{d\in\mathbb{N}}$ of the r.h.s., and by the very definition of lim inf we get:

$$\limsup_{b\to\infty} \frac{a_b}{b} \leq \inf_{d\in\mathbb{N}} \frac{a_d}{d} \leq \liminf_{d\to\infty} \frac{a_d}{d} \tag{1.7}$$

The latter implies that lim inf and lim sup are equal and consequently the existence of the limit. Finally, (1.4) implies that the sequence must approach its $\inf_{\mathbb{N}}$. Hence:

$$\lim_{n\to\infty} \frac{a_n}{n} = \inf_{b\in\mathbb{N}} \frac{a_b}{b} \tag{1.8}$$

The result is valid also with a super-additivity hypothesis, provided that each inequality is reversed in the proof and that the $\inf_{\mathbb{N}}$ in (1.2) is replaced with a $\sup \mathbb{N}$. □

Remark 1.1.1. The sequence we are interested in is the pressure P_N. The hypothesis $N\underline{c} \leq P_N \leq N\bar{c}$ is rather reasonable. The linear growth of the pressure with the number of particles in the system is a desirable feature for any well posed model. In fact, pressure and Hamiltonian must be extensive quantities, and grow roughly linearly with the size of the system in order to have the usual physical interpretation.

1.1.2 Existence through symmetrization lemma

The first method to prove the existence of the limit, which we show here, makes use of the so called *Symetrization* or *Urns lemma*.

It is important to stress that the following results are applicable only to a model whose Hamiltonian is:

$$H_N^{CW}(\sigma) = \frac{-J}{N-1} \sum_{i,j\in I_N,\, i<j} \sigma_i\sigma_j - h \sum_{i\in I_N} \sigma_i \tag{1.9}$$

where $I_N = \{1, 2, \ldots, N\}$. To begin with, we prove the following proposition.

Proposition 1.1.2 (Lower and Upper bound for the pressure per particle)**.** *Given the pressure of the model* (1.9) *with* $J > 0$*:*

$$P_N^{CW} = \log \left[\sum_{\sigma\in\Sigma_N} \exp\left(-\beta H_N^{CW}(\sigma)\right) \right] \tag{1.10}$$

there are two real constants $\underline{c}, \bar{c} \in \mathbb{R} : N\underline{c} \leq P_N \leq N\bar{c}$*.*

Proof. Let us focus on the exponential Boltzmann factor:

$$\exp\left(-\beta H_N^{CW}(\sigma)\right) \le \exp\left(-\beta \min_{\sigma\in\Sigma_N} H_N^{CW}(\sigma)\right) = \exp\left(-\beta H_N^{CW}(\sigma^*)\right) \tag{1.11}$$

$\sigma^* \in \Sigma_N$ is a configuration that minimizes the Hamiltonian.

Since $\exp\left(-\beta H_N^{CW}(\sigma^*)\right)$ is a term of the sum in (1.10), we get:

$$\sum_{\sigma\in\Sigma_N} \exp\left(-\beta H_N^{CW}(\sigma)\right) \le \sum_{\sigma\in\Sigma_N} \exp\left(-\beta H_N^{CW}(\sigma^*)\right) = 2^N \exp\left(-\beta H_N^{CW}(\sigma^*)\right) \tag{1.12}$$

$$\sum_{\sigma\in\Sigma_N} \exp\left(-\beta H_N^{CW}(\sigma)\right) \ge \exp\left(-\beta H_N^{CW}(\sigma^*)\right) \tag{1.13}$$

Gathering the two inequalities and taking the logarithm we get:

$$\frac{\beta J}{N-1}\sum_{i,j\in I_N,\, i<j} \sigma_i^*\sigma_j^* + \beta h \sum_{i\in I_N}\sigma_i^* \le P_N \le N\log 2 + \frac{\beta J}{N-1}\sum_{i,j\in I_N,\, i<j} \sigma_i^*\sigma_j^* + \beta h \sum_{i\in I_N}\sigma_i^* \tag{1.14}$$

In the ferromagnetic case, *i.e.* $J>0$, there is only one configuration that minimizes the Hamiltonian: $\sigma_i^* = sing(h)$ for all $i \in I_N$. Inserting this spin configuration:

$$N(\beta J + \beta|h|) \le P_N \le N(\log 2 + \beta J + \beta|h|) \tag{1.15}$$

□

Lower and upper bounds can also be immediately found by a quick inspection on the hamiltonian, as stated below.

Proposition 1.1.3 (Lower (upper) bound for the generating functional P)**.** *Consider the generating functional:*

$$P_N = \log \mathbb{E}_\sigma \exp\left[-\beta H_N(\sigma)\right] \tag{1.16}$$

where $H_N(\sigma)$ depends on a collection of random variables $(\sigma_i)_{1\le i\le N}$ and $H_N(\sigma) \le KN$ (respectively $H_N(\sigma) \ge KN$) $\forall N \in \mathbb{N}$ and for a certain $K \in \mathbb{R}$.

The sequence P_N/N is bounded from below (above), i.e. *$\exists\, \underline{c}$ (or $\bar{c}$) $\in \mathbb{R} : P_N/N \ge \underline{c}$ ($P_N/N \le \bar{c}$) $\forall N \in \mathbb{N}$.*

Proof. Since $\exp(-x)$ is a monotonically decreasing function we have:

$$P_N = \log \mathbb{E}_\sigma \exp\left[-\beta H_N(\sigma)\right] \ge \log \mathbb{E}_\sigma \exp\left[-\beta K N\right] \tag{1.17}$$

The exponential now is independent on the random variables, the expectation is thus trivial.

$$P_N \geq -K\beta N \quad \Rightarrow \quad \underline{c} = -\beta K \tag{1.18}$$

K can be either positive or negative, in both cases its sign does not affect the result. The proof of the reversed inequality follows the same steps. □

Remark 1.1.2. The requirement that H_N has an extensive growth in N is reasonable as pointed out previously.

Furthermore, notice that, even with the addition of a logarithm as shown here:

$$H_N(\sigma) \leq KN + C\log N \tag{1.19}$$

the hypothesis are still verified. Leaving aside the trivial case $C = 0$, for $C > 0$ we have:

$$\frac{P_N}{N} \geq -\beta K - \beta C\frac{N-1}{N} \geq -\beta(K+C) = \underline{c} \tag{1.20}$$

Let us discuss $C < 0$. $\log N/N$ is always non negative, because $N \geq 1$. Hence:

$$\frac{P_N}{N} \geq -\beta K + \beta|C|\frac{\log N}{N} \geq -\beta K = \underline{c} \tag{1.21}$$

In order to formulate the symmetrization lemma we need to define the following tools.

Definition 1.1.1 (Bipartition of a set of indices)**.** Let $I_N = \{1, 2, 3, ..., N\}$ be a set of N indices. A couple (I_{N_1}, I_{N_2}) of disjoint subsets of I_N such that $N_1 + N_2 = N$, $N_1, N_2 \geq 2$ and $I_{N_1} \uplus I_{N_2} = I_N$ is called *Bipartition* of I_N. Furthermore I_{N_1} and I_{N_2} are called *left* and *right partition* respecively.

The reader has certainly noticed that, once the cardinalities, or the sizes of the urns, N_1 and N_2 are fixed, there are many ways to divide N distinguishable indices into the two urns. Thus we need to introduce a family of possible bipartitions.

Notation 1.1.1 (Family of bipartitions of I_N)**.** Given a set of indices I_N, the set of its possible bipartitions in N_1 and N_2 elements will be denoted by:

$$\mathcal{P}_{N_1,N_2} = \{(I_{N_1}, I_{N_2})|(I_{N_1}, I_{N_2}) \text{ is a bipartition of } I_N,\ |I_{N_1}| = N_1,\ |I_{N_2}| = N_2\} \tag{1.22}$$

Remark 1.1.3. The cardinality of the set $\mathcal{P}_{N_1,N_2}$ can be easily computed in a combinatorial way. In fact, one can always consider it as the number of possible ways to extract N_1 indices out of a set of N distinguishable indices. The order of extraction is irrelevant. Notice that after this extraction the set of the remaining N_2 indices is automatically determined. We finally get:

$$|\mathcal{P}_{N_1,N_2}| = \mathbf{C}^N_{N_1} = \frac{N!}{N_1!(N-N_1)!} = \frac{N!}{N_1!N_2!} \tag{1.23}$$

We are now ready to write and prove the statement.

Lemma 1.1.4 (Symmetrization)**.** *Let $I_N = \{1, 2, ..., N\}$ be a set of indices and f_{ij}, with $i, j \in I_N$, $i \neq j$, a collection of real numbers. Consider a function of these numbers defined as:*

$$H_N(f) = \frac{1}{N-1} \sum_{i,j \in I_N,\, i \neq j} f_{ij} \tag{1.24}$$

H_N can always be written in the following form:

$$H_N(f) = \frac{1}{|\mathcal{P}_{N_1,N_2}|} \sum_{(I_{N_1}, I_{N_2}) \in \mathcal{P}_{N_1,N_2}} [H_{N_1}(f) + H_{N_2}(f)] \tag{1.25}$$

where H_{N_1} and H_{N_2} are the restrictions of the function to the left and right partitions respectively:

$$H_{N_1}(f) = \frac{1}{N_1 - 1} \sum_{i,j \in I_{N_1},\, i \neq j} f_{ij} \quad ; \quad H_{N_2}(f) = \frac{1}{N_2 - 1} \sum_{i,j \in I_{N_2},\, i \neq j} f_{ij} \tag{1.26}$$

Proof. For fixed $i, j \in I_N$, $i \neq j$, at most one of the two restrictions of the function H_N contains the term f_{ij}, because the two partitions are disjoint. It is easy to realize that while computing the sum over the bipartitions, each f_{ij} appears more than once in the left and right partitions. The contribution it gives to the total sum is equal to:

$$\frac{N_1!N_2!}{N!} f_{ij} \left[\frac{1}{N_1 - 1} \frac{(N-2)!}{(N_1-2)!(N-N_1)!} + \frac{1}{N_2 - 1} \frac{(N-2)!}{(N_2-2)!(N-N_2)!} \right] \tag{1.27}$$

Supposing $i, j \in I_{N_1}$, the first term is nothing but the number of ways of picking $N_1 - 2$ indices out of $N-2$, divided by the normalization inherited from the restriction. The second term is computed in the same way, but i, j are fixed in the right partition I_{N_2} this time.

Let us focus on the square bracket.

$$\frac{1}{N_1-1}\frac{(N-2)!}{(N_1-2)!(N-N_1)!}+\frac{1}{N_2-1}\frac{(N-2)!}{(N_2-2)!(N-N_2)!}=$$
$$=\frac{(N-2)!}{(N_1-1)!N_2!}+\frac{(N-2)!}{(N_2-1)!N_1!}=\frac{(N-2)!}{N_1!N_2!}\left[N_1+N_2\right]=\frac{1}{N-1}\frac{N!}{N_1!N_2!} \tag{1.28}$$

Notice that the last combinatorial factor is exactly the cardinality of $\mathcal{P}_{N_1,N_2}$.

We have just shown that the contribution each f_{ij} gives to the total sum is:

$$\frac{N_1!N_2!}{N!}f_{ij}\left[\frac{1}{N-1}\frac{N!}{N_1!N_2!}\right]=\frac{f_{ij}}{N-1} \tag{1.29}$$

This proves the claim. □

At a first sight, this result does not seem to have an immediate interpretation. However, besides the technicalities, the statement is quite simple: one can always construct a function of the type (1.24) if every restriction of it to all possible partitions of the set of indices, of fixed cardinality, is known.

This is rather intuitive, and can also acquire a physical meaning when dealing with Hamiltonians. In this case the two Hamiltonians restricted to the left and right partitions represent *two non-interacting Curie-Weiss models*, whose total Hamiltonian is exactly the sum of them. This observation will be crucial to prove the existence of the thermodynamic limit. Therefore this lemma tells us that we can study the thermodynamic properties of a Curie-Weiss system through non interacting subsystems of it. The interactions are re-introduced by the sum over all the possible bipartitions, as the proof of the lemma suggests.

Remark 1.1.4. The Hamiltonian of the CW model (1.9) is a function of the type (1.24) because the introduction of a one-body term does not alter the hypothesis of the Lemma. To show this, let us consider a set of real numbers g_i, $i \in I_N$ and the new Hamiltonian:

$$H_N(f,g)=\frac{1}{N-1}\sum_{i,j\in I_N,\, i\neq j} f_{ij}+\sum_{i\in I_N} g_i \tag{1.30}$$

Observe that, since for each fixed i there are $N-1$ js different from i, we have:

$$\sum_{i\in I_N} g_i=\frac{1}{N-1}\sum_{i,j\in I_N,\, i\neq j} g_i \tag{1.31}$$

$$H_N(f,g)=\frac{1}{N-1}\sum_{i,j\in I_N,\, i\neq j} f_{ij}+\sum_{i\in I_N} g_i=\frac{1}{N-1}\sum_{i,j\in I_N,\, i\neq j}(f_{ij}+g_i) \tag{1.32}$$

The previous Hamiltonian then reduces to (1.9) with the identifications: $f_{ij} = -(J/2)\sigma_i\sigma_j$, $g_i = h\sigma_i$.

An analogous result is valid for a Hamiltonian with k-body terms, as shown below.

Lemma 1.1.5 (Symmetrization lemma, k-body terms)**.** *Let $I_N = \{1, 2, ..., N\}$ be a set of indices and $f_{i_1 i_2 \dots i_k}$, with $i_1, i_2, ..., i_k \in I_N$, $k \leq N$ and $i_l \neq i_j \, \forall l, j \in \{1, 2, ..., k\}$. The left and right partitions must contain $N_1, N_2 \geq k$ indices. Then:*

$$H_N(f) = \frac{(N-k)!}{(N-1)!} \sum_{(i_1, i_2, \dots, i_k) \in I_N} f_{i_1 i_2 \dots i_k} = \frac{1}{|\mathcal{P}_{N_1,N_2}|} \sum_{(I_{N_1}, I_{N_2}) \in \mathcal{P}_{N_1,N_2}} [H_{N_1}(f) + H_{N_2}(f)] \tag{1.33}$$

where $(i_1, ..., i_k)$ indicates that the sum is computed over different indices, and:

$$H_{N_1}(f) = \frac{(N_1-k)!}{(N_1-1)!} \sum_{(i_1, i_2, \dots, i_k) \in I_{N_1}} f_{i_1 i_2 \dots i_k} \quad ; \quad H_{N_2}(f) = \frac{(N_2-k)!}{(N_2-1)!} \sum_{(i_1, i_2, \dots, i_k) \in I_{N_2}} f_{i_1 i_2 \dots i_k} \tag{1.34}$$

Proof. The proof follows the same strategy used for the two-body version of this lemma. For fixed $(i_1, ..., i_k)$ and N_1, $N_2 \geq k$, the contribution of each $f_{i_1,\dots,i_k}$ is:

$$f_{i_1,\dots,i_k} \left[\frac{(N_1-k)!}{(N_1-1)!} \frac{(N-k)!}{(N-N_1)!(N_1-k)!} + \frac{(N_2-k)!}{(N_2-1)!} \frac{(N-k)!}{(N-N_2)!(N_2-k)!} \right] \tag{1.35}$$

Again, the first factor in the first term is the normalization inherited from the restriction H_{N_1}, while the second factor is the number of ways of picking $N_1 - k$ indices out of $N - k$ indices (k indices are already fixed in the left partition). The second term, analogously, takes into account the numbers of times $f_{i_1,\dots,i_k}$ appears in the right partition.

The square bracket can be simplified:

$$\begin{aligned} \frac{(N-k)!}{(N-N_1)!(N_1-1)!} + \frac{(N-k)!}{(N-N_2)!(N_2-1)!} &= \frac{(N-k)!}{N_1!N_2!}[N1 + N_2] = \\ &= \frac{N(N-k)!}{N_1!N_2!} = \frac{N!}{N_1!N_2!} \frac{(N-k)!}{(N-1)!} \end{aligned} \tag{1.36}$$

After including the cardinality of the bipartitions family we get:

$$\frac{N_1!N_2!}{N!} f_{i_1,\dots,i_k} \frac{N!}{N_1!N_2!} \frac{(N-k)!}{(N-1)!} = f_{i_1,\dots,i_k} \frac{(N-k)!}{(N-1)!} \tag{1.37}$$

□

Remark 1.1.5. It is now easy to see that the previous lemma reduces to the two-body case when $k = 2$. Moreover, one can show that the introduction of a $(k-1)$-body term does not alter the form of the function, from the point of view of the lemma. To see this let us consider the function:

$$H_N(f) = \frac{(N-k)!}{(N-1)!} \sum_{(i_1,i_2,\dots,i_k)\in I_N} f_{i_1 i_2 \dots i_k} + \frac{(N-k+1)!}{(N-1)!} \sum_{(i_1,i_2,\dots,i_{k-1})\in I_N} g_{i_1 i_2 \dots i_{k-1}} \tag{1.38}$$

Notice that, for fixed $(i_1, \dots, i_{k-1})$ there are $N-(k-1)$ possible values for i_k, then:

$$\begin{aligned}\frac{(N-k)!}{(N-1)!} \sum_{(i_1,i_2,\dots,i_k)\in I_N} g_{i_1 i_2 \dots i_{k-1}} &= \frac{(N-k)!(N-k+1)}{(N-1)!} \sum_{(i_1,i_2,\dots,i_{k-1})\in I_N} g_{i_1 i_2 \dots i_{k-1}} = \\ &= \frac{(N-k+1)!}{(N-1)!} \sum_{(i_1,i_2,\dots,i_{k-1})\in I_N} g_{i_1 i_2 \dots i_{k-1}} \end{aligned} \tag{1.39}$$

Thus, we can rewrite the function H_N as follows:

$$H_N(f) = \frac{(N-k)!}{(N-1)!} \sum_{(i_1,i_2,\dots,i_k)\in I_N} [f_{i_1 i_2 \dots i_k} + g_{i_1 i_2 \dots i_{k-1}}] = \frac{(N-k)!}{(N-1)!} \sum_{(i_1,i_2,\dots,i_k)\in I_N} \tilde{f}_{i_1 i_2 \dots i_k} \tag{1.40}$$

and the lemma can still be applied.

Notation 1.1.2. Let me introduce this useful notation:

$$\mathcal{E}_{N_1,N_2}\left[H_{N_1}(f) + H_{N_2}(f)\right] = \frac{1}{|\mathcal{P}_{N_1,N_2}|} \sum_{(I_{N_1},I_{N_2})\in\mathcal{P}_{N_1,N_2}} [H_{N_1}(f) + H_{N_2}(f)] \tag{1.41}$$

Remark 1.1.6. The use of letter $\mathcal{E}$, that stands for "expectation", is justified by the fact that the r.h.s. of (1.41) can be considered as an expectation computed according to a uniform probability distribution defined on the family $\mathcal{P}_{N_1,N_2}$.

We have developed all the necessary mathematical tools to prove the following theorem.

Theorem 1.1.6 (Sub-additivity of the generating functional P_N)**.** *Let $(\sigma_i)_{i\in I_N}$ be a set of i.i.d. random variables and g_{ij}, with $i,j \in I_N$, $i \neq j$, a set of functions of (σ_i, σ_j) only. Then, denoting by $\mathbb{E}_\sigma$ the expectation on the r.v., the generating functional:*

$$P_N = \log \mathbb{E}_\sigma \exp\left[-\beta H_N(\sigma)\right] \quad ; \quad H_N(\sigma) = \frac{1}{N-1} \sum_{i,j\in I_N,\, i\neq j} g_{ij}(\sigma_i, \sigma_j) \tag{1.42}$$

is sub-additive i.e. $P_N \leq P_{N_1} + P_{N_2}$

Proof.

$$P_N = \log \mathbb{E}_\sigma \exp\left[-\beta H_N(\sigma)\right] = \log \mathbb{E}_\sigma \exp\left[-\beta \mathcal{E}_{N_1,N_2}(H_{N_1}(\sigma) + H_{N_2}(\sigma))\right] \quad (1.43)$$

Using Jensen's inequality we have:

$$\begin{aligned} P_N \le \log \mathbb{E}_\sigma \mathcal{E}_{N_1,N_2} \exp\left[-\beta(H_{N_1}(\sigma) + H_{N_2}(\sigma))\right] = \\ = \log \mathcal{E}_{N_1,N_2} \mathbb{E}_\sigma \exp\left[-\beta(H_{N_1}(\sigma) + H_{N_2}(\sigma))\right] \end{aligned} \quad (1.44)$$

Since the two partitions I_{N_1} and I_{N_1} are disjoint, if H_{N_1} depends on a variable σ_i then H_{N_2} cannot depend on the same variable. This means that the expectations splits into the product of two expectations:

$$P_N \le \log\left[\mathcal{E}_{N_1,N_2} \mathbb{E}_\sigma \exp(-\beta H_{N_1}(\sigma)) \mathbb{E}_\sigma \exp(-\beta H_{N_2}(\sigma))\right] \quad (1.45)$$

Once the expectation has been computed, the result is a constant to the uniform measure on the bipartitions family, because the random variables are i.i.d., hence:

$$P_N \le \log\left[\mathbb{E}_\sigma \exp(-\beta H_{N_1}(\sigma)) \mathbb{E}_\sigma \exp(-\beta H_{N_2}(\sigma))\right] = P_{N_1} + P_{N_2} \quad (1.46)$$

□

Corollary 1.1.7 (Sub-additivity of Curie-Weiss pressure P_N)**.** *Given the Hamiltonian:*

$$H_N^{CW}(\sigma) = \frac{-J}{N-1} \sum_{i,j \in I_N,\, i<j} \sigma_i \sigma_j - h \sum_{i \in I_N} \sigma_i \quad (1.47)$$

then the corresponding pressure is sub-additive, i.e.

$$P_N(\beta; J, h) = P_{N_1+N_2}(\beta; J, h) \le P_{N_1}(\beta; J, h) + P_{N_1}(\beta; J, h) \quad (1.48)$$

Proof. We have already shown that (1.9) is of the type (1.24), even with the addition of a one-body term. Furthermore we have precisely the hypothesis of the previous theorem.

$$P_N = \log\left[2^N \omega_0\left(\exp(-\beta H_N^{CW}(\sigma)\right)\right] = N \log 2 + \log\left[\omega_0\left(\exp(-\beta H_N^{CW}(\sigma)\right)\right] \quad (1.49)$$

where the expectation is taken with respect to the uniform distribution:

$$\omega_0[\cdot] = \frac{1}{2^N} \sum_{\sigma \in \Sigma_N} [\cdot] \quad (1.50)$$

which is a product distribution. Hence

$$\begin{aligned} P_N = N \log 2 + \log\left[\omega_0\left(\exp(-\beta H_N^{CW}(\sigma)\right)\right] \le (N_1 + N_2) \log 2 + \\ + \log\left[\omega_0\left(\exp(-\beta H_{N_1}^{CW}(\sigma)\right)\right] + \log\left[\omega_0\left(\exp(-\beta H_{N_2}^{CW}(\sigma)\right)\right] = P_{N_1} + P_{N_2} \end{aligned} \quad (1.51)$$

□

1.1.3 Existence through Guerra-Toninelli's interpolation

From this moment on we will work with the Hamiltonian:

$$H_N^{CW}(\sigma) = \frac{-J}{N} \sum_{i,j \in I_N} \sigma_i \sigma_j - h \sum_{i \in I_N} \sigma_i \tag{1.52}$$

where $I_N = \{1, 2, ..., N\}$. The diagonal terms in the first sum do not affect the thermodynamics in the limit of large N. In order to simplify the form of the Hamiltonian we introduce the configurational magnetization:

Definition 1.1.2 (Configurational magnetization)**.** It is a number $-1 \leq m_N(\sigma) \leq +1$, defined as:

$$m_N(\sigma) = \frac{1}{N} \sum_{i \in I_N} \sigma_i \tag{1.53}$$

(1.52) then becomes:

$$H_N^{CW}(\sigma) = -JNm_N^2(\sigma) - hNm_N(\sigma) \tag{1.54}$$

We see that the Hamiltonian has a linear growth in N as it should be.

Here the idea is to interpolate continuously the system with N spins with two separated, non-interacting subsystems, by means of the interpolating Hamiltonian.

Definition 1.1.3 (Interpolating Hamiltonian and pressure)**.** The interpolating Hamiltonian is a convex combination of the Hamiltonians H_N^{CW} and $H_{N_1}^{CW} + H_{N_2}^{CW}$:

$$H(t) = tH_N^{CW}(\sigma) + (1-t)[H_{N_1}^{CW}(\sigma) + H_{N_2}^{CW}(\sigma)] \quad t \in [0,1] \tag{1.55}$$

The interpolating pressure is:

$$P(t) = \log Z(t) \quad \text{with} \quad Z(t) = \sum_{\sigma \in \Sigma_N} \exp\left(-\beta H(t)\right) \tag{1.56}$$

Remark 1.1.7.

$$P(1) = \log\left[\sum_{\sigma \in \Sigma_N} \exp\left(-\beta H(1)\right)\right] = \log\left[\sum_{\sigma \in \Sigma_N} \exp\left(-\beta H_N^{CW}(\sigma)\right)\right] \tag{1.57}$$

$$P(0) = \log\{\sum_{\sigma \in \Sigma_N} \exp\left[-\beta(H_{N_1}^{CW}(\sigma) + H_{N_2}^{CW}(\sigma))\right]\} = P_{N_1} + P_{N_2} \tag{1.58}$$

Proving sub-additivity is thus equivalent to prove that: $P(1) \leq P(0)$.

Theorem 1.1.8 (Sub-additivity of the pressure through interpolation)**.** *The interpolating pressure $P(t)$ of a ferromagnetic Curie-Weiss model is monotonically decreasing in $t \in [0,1]$, and $P(1) \leq P(0)$.*

Proof. We calculate the first derivative with respect to t of $P(t)$.

$$P'(t) = \frac{-\beta}{Z(t)} \sum_{\sigma \in \Sigma_N} e^{-\beta H(t)}[H_N^{CW}(\sigma) - (H_{N_1}^{CW}(\sigma) + H_{N_2}^{CW}(\sigma))] = \\ = -\beta \omega_{N,t}(H_N^{CW}(\sigma) - H_{N_1}^{CW}(\sigma) - H_{N_2}^{CW}(\sigma)) \quad (1.59)$$

$\omega_{N,t}(\cdot)$ denotes the expectation with respect to the Gibbs measure.

$$H_N^{CW}(\sigma) - H_{N_1}^{CW}(\sigma) - H_{N_2}^{CW}(\sigma) = -J[N m_N^2(\sigma) - N_1 m_{N_1}^2(\sigma) - N_2 m_{N_2}^2(\sigma)] - \\ - h[N m_N(\sigma) - N_1 m_{N_1}(\sigma) - N_2 m_{N_2}(\sigma)] \quad (1.60)$$

The last term vanishes because of the definition of configurational magnetization. We use this fact to substitute $m_N(\sigma)$ in the first term.

$$- JN\left[m_N^2(\sigma) - \frac{N_1}{N} m_{N_1}^2(\sigma) - \frac{N_2}{N} m_{N_2}^2(\sigma)\right] = \\ = -JN\left[\left(\frac{N_1}{N} m_{N_1}(\sigma) + \frac{N_2}{N} m_{N_2}(\sigma)\right)^2 - \frac{N_1}{N} m_{N_1}^2(\sigma) - \frac{N_2}{N} m_{N_2}^2(\sigma)\right] \geq 0 \quad (1.61)$$

The last inequality follows from the convexity of the function $f(x) = x^2$. Finally, the positivity of the expectation functional leads us to:

$$P'(t) = -\beta \omega_{N,t}(H_N^{CW}(\sigma) - H_{N_1}^{CW}(\sigma) - H_{N_2}^{CW}(\sigma)) \leq 0 \quad \Rightarrow \quad P(1) \leq P(0) \quad (1.62)$$

□

Remark 1.1.8. In the case of an antiferromagnetic model, *i.e.* $J < 0$ the sign of $P'(t)$ would change and this would lead to a super-additive pressure. The limit exists anyway thanks to the Lemma 1.1.4, but the sequence P_N/N converges towards its $\sup_{\mathbb{N}}$.

Thanks to the interpolation technique we can also finally prove that a slight change in the normalization of the two body term does not affect the thermodynamic limit. This can be done both for a deterministic model, such as Curie-Weiss, and for a disordered system in which interactions are random, such as Sherrington-Kirkpartick.

Theorem 1.1.9 (Normalization stability of the thermodynamic limit, CW). *Consider the two Hamiltonians:*

$$H_N(\sigma) = \frac{-J}{N} \sum_{i,j \in I_N, i<j} \sigma_i \sigma_j - h \sum_{i \in I_N} \sigma_i \tag{1.63}$$

$$\tilde{H}_N(\sigma) = \frac{-J}{N+c} \sum_{i,j \in I_N, i<j} \sigma_i \sigma_j - h \sum_{i \in I_N} \sigma_i \tag{1.64}$$

where $c \in \mathbb{R}$. The two Hamiltonians induce the same pressure per particle in the thermodynamic limit:

$$\lim_{N \to \infty} (p_N - \tilde{p}_N) = 0 \tag{1.65}$$

$$p_N = \frac{1}{N} \log \sum_{\sigma \in \Sigma_N} e^{-\beta H_N(\sigma)} \tag{1.66}$$

$$\tilde{p}_N = \frac{1}{N} \log \sum_{\sigma \in \Sigma_N} e^{-\beta \tilde{H}_N(\sigma)} \tag{1.67}$$

Proof. In order to compare the two Hamiltonians we can interpolate them with a parameter $t \in [0,1]$. Then the computation of the interpolating pressure is straightforward.

$$H(t) = \frac{-J}{N+tc} \sum_{i,j \in I_N, i<j} \sigma_i \sigma_j - h \sum_{i \in I_N} \sigma_i \tag{1.68}$$

$$P(t) = \log \sum_{\sigma \in \Sigma_N} e^{-\beta H(t)} \tag{1.69}$$

$$P(0) = P_N = \log \sum_{\sigma \in \Sigma_N} e^{-\beta H_N(\sigma)} \quad ; \quad P(1) = \tilde{P}_N = \log \sum_{\sigma \in \Sigma_N} e^{-\beta \tilde{H}_N(\sigma)} \tag{1.70}$$

The difference between the two pressures can be computed through the first derivative with respect to t of $P(t)$, that will play an important role.

$$P(t) = \log \sum_{\sigma \in \Sigma_N} \exp \left(\frac{\beta J}{N+tc} \sum_{i,j \in I_N, i<j} \sigma_i \sigma_j + \beta h \sum_{i \in I_N} \sigma_i \right) \tag{1.71}$$

$$\frac{d}{dt} P(t) = \beta \omega_{N,t} \left[\frac{-Jc}{(N+tc)^2} \sum_{i,j \in I_N, i<j} \sigma_i \sigma_j \right] = \frac{-\beta J c}{(N+tc)^2} \sum_{i,j \in I_N, i<j} \omega_{N,t}[\sigma_i \sigma_j] \tag{1.72}$$

The expectation is taken with respect to the Gibbs measure induced by the interpolating Hamiltonian. This measure is invariant under permutations of the spins, thanks to the form of $H(t)$. Using this fact, one realizes that the $N(N-1)/2$ terms in the sum are all equal.

$$\frac{d}{dt}P(t) = \frac{-\beta JcN(N-1)}{2(N+tc)^2}\omega_{N,t}[\sigma_1\sigma_2] \tag{1.73}$$

Finally, by means of $P'(t)$ we are able to estimate the difference $p_N - \tilde{p}_N$.

$$\tilde{p}_N - p_N = p(1) - p(0) = \frac{1}{N}\int_0^1 dt \frac{d}{dt}P(t) = \frac{-\beta Jc(N-1)}{2}\int_0^1 dt \frac{\omega_{N,t}[\sigma_1\sigma_2]}{(N+tc)^2} \tag{1.74}$$

Let us focus on the quantity:

$$\left|(N-1)\int_0^1 dt \frac{\omega_{N,t}[\sigma_1\sigma_2]}{(N+tc)^2}\right| \leq (N-1)\int_0^1 dt \frac{1}{(N+tc)^2} = \frac{N-1}{N(N+c)} \longrightarrow 0 \tag{1.75}$$

when $N \to \infty$. Hence:

$$|p_N - \tilde{p}_N| \longrightarrow 0 \tag{1.76}$$

□

1.1.4 Alternative proof of sub-additivity

To conclude this section, we provide an alternative proof of the sub-additivity of the pressure that works with both normalizations N and $N-1$ seen before. This proof is due to Francesco Guerra, but it can also be found in [8].

Lemma 1.1.10 (Alternative proof of sub-additivity)**.** *Consider a system with Hamiltonian H_N, and divide it into two subsystems with Hamiltonians H_{N_1} and H_{N_2}. If the following condition holds:*

$$\omega_N[H_N - H_{N_1} - H_{N_2}] \geq 0 \quad \textit{with} \quad \omega_N[\cdot] = \sum_{\sigma\in\Sigma_N}[\cdot]\frac{e^{-\beta H_N(\sigma)}}{\sum_{\tau\in\Sigma_N} e^{-\beta H_N(\tau)}} \tag{1.77}$$

the pressure is sub-additive.

Proof. Let us compare the partition function of the the global system, whose Hamiltonian is H_N, to that of the composed system, whose Hamiltonian is $H_{N_1} + H_{N_2}$. We proceed as in the following:

$$\frac{Z_{N_1}Z_{N_2}}{Z_N} = \frac{\sum_{\sigma\in\Sigma_N} e^{-\beta H_{N_1}(\sigma)-\beta H_{N_2}(\sigma)}}{\sum_{\tau\in\Sigma_N} e^{-\beta H_N(\tau)}} = \sum_{\sigma\in\Sigma_N} e^{\beta(H_N(\sigma)-H_{N_1}(\sigma)-H_{N_2}(\sigma))}\frac{e^{-\beta H_N(\sigma)}}{Z_N} \tag{1.78}$$

We have implicitly used the fact that the partition function a system composed by two non interacting subsystems is the product of the partition functions of the latter. We can recognize a Gibbs measure in the last factor and this makes it possible to use Jensen's inequality:

$$\frac{Z_{N_1}Z_{N_2}}{Z_N} = \omega_N\left[e^{\beta(H_N(\sigma)-H_{N_1}(\sigma)-H_{N_2}(\sigma))}\right] \geq e^{\beta\omega_N\left[H_N(\sigma)-H_{N_1}(\sigma)-H_{N_2}(\sigma)\right]} \geq 1 \quad (1.79)$$

The last inequality directly follows from the hypothesis. The partition function is sub-multiplicative, hence the pressure is sub-additive. □

Lemma 1.1.11. *The Hamiltonian* (1.52) *satisfies the hypothesis of Lemma 1.1.10.*

Proof. The proof is contained in the proof of Theorem 1.1.8. In fact:

$$H_N - H_{N_1} - H_{N_2} = -JN\left[m_N^2(\sigma) - \frac{N_1}{N}m_{N_1}^2(\sigma) - \frac{N_2}{N}m_{N_2}^2(\sigma)\right] =$$
$$= -JN\left[\left(\frac{N_1}{N}m_{N_1}(\sigma) + \frac{N_2}{N}m_{N_2}(\sigma)\right)^2 - \frac{N_1}{N}m_{N_1}^2(\sigma) - \frac{N_2}{N}m_{N_2}^2(\sigma)\right] \geq 0 \quad (1.80)$$

The expectation is a positive functional, hence $\omega_N[H_N - H_{N_1} - H_{N_2}] \geq 0$. □

Lemma 1.1.12. *The Hamiltonian* (1.9) *satisfies the hypothesis of Lemma 1.1.10.*

Proof. Let us focus on the single expectations:

$$\omega_N[H_N] = \omega_N\left[\frac{-J}{N-1}\sum_{i<j}\sigma_i\sigma_j - h\sum_i\sigma_i\right] = \frac{-J}{N-1}\sum_{i<j}\omega_N\left[\sigma_i\sigma_j\right] - h\sum_i\omega_N[\sigma_i] \quad (1.81)$$

Now we can use the fact the the measure in (1.77) enjoys invariance under permutation of the spins, so the expectations $\omega_N[\sigma_i\sigma_j] = \omega_N[\sigma_1\sigma_2]$, $i \neq j$ and $\omega_N[\sigma_i] = \omega_N[\sigma_1]$. The previous formula becomes:

$$\omega_N[H_N] = \frac{-J\omega_N[\sigma_1\sigma_2]}{N-1}\frac{N(N-1)}{2} - hN\omega_N[\sigma_1] = -\frac{J}{2}N\omega_N[\sigma_1\sigma_2] - hN\omega_N[\sigma_1] \quad (1.82)$$

Analogous results are valid for $\mathbb{E}[H_{N_1}]$ and $\mathbb{E}[H_{N_2}]$.

$$\omega_N[H_{N_1}] = -\frac{J}{2}N_1\omega_N[\sigma_1\sigma_2] - hN_1\omega_N[\sigma_1] \quad (1.83)$$
$$\omega_N[H_{N_2}] = -\frac{J}{2}N_2\omega_N[\sigma_1\sigma_2] - hN_2\omega_N[\sigma_1] \quad (1.84)$$

We finally get:

$$\omega_N[H_N - H_{N_1} - H_{N_2}] = -\frac{J}{2}(N - N_1 - N_2)\omega_N[\sigma_1\sigma_2] - h(N - N_1 - N_2)\omega_N[\sigma_1] = 0 \tag{1.85}$$

□

1.2 Normalization effects

1.2.1 Pressure and ground state energy

Definition 1.2.1 (Ground state energy per particle)**.** Given an Hamiltonian $H_N(\sigma; J, h)$, where $\sigma \in \Sigma_N$ is a generic spin configuration, the ground state energy per particle is defined as:

$$e_N(J, h) = \inf_{\sigma \in \Sigma_N} \frac{H_N(\sigma; J, h)}{N} \tag{1.86}$$

This definition is still valid for disordered systems, where the couplings J are random.

Proposition 1.2.1 (Low temperature limit of the pressure per particle)**.** *Given the pressure per particle $p_N(\beta; J, h)$ induced by an Hamiltonian $H(\sigma; J, h)$ we have:*

$$\lim_{\beta \to \infty} \frac{p_N(\beta; J, h)}{\beta} = -e_N(J, h) \tag{1.87}$$

Proof. First we seek for upper and lower bounds for the partition function.

$$\begin{aligned} \exp\left(-\beta N \inf_{\sigma \in \Sigma_N} \frac{H_N(\sigma; J, h)}{N}\right) \leq Z_N(\beta; J, h) = \sum_{\sigma \in \Sigma_N} e^{-\beta H_N(\sigma; J, h))} \leq \\ \leq 2^N \exp\left(-\beta N \inf_{\sigma \in \Sigma_N} \frac{H_N(\sigma; J, h)}{N}\right) \end{aligned} \tag{1.88}$$

After that we take the logarithm divided by βN.

$$-e_N(J, h) \leq \frac{p_N(\beta; J, h)}{\beta} \leq -e_N(J, h) + \frac{\log 2}{\beta} \tag{1.89}$$

Both the lower and the upper bounds approach $-e_N(J, h)$ when $\beta \to \infty$. This proves the claim. □

Proposition 1.2.2 (Monotonicity of the large N limit of $e_N(J,h)$)**.** *Consider a deterministic ferromagnetic model whose Hamiltonian is:*

$$H_N(\sigma; J, h) = -\frac{J}{N+c} \sum_{i,j \in I_N,\, i<j} \sigma_i \sigma_j - h \sum_{i \in I_N} \sigma_i \quad \textit{with } J > 0,\ c \in \mathbb{R} \tag{1.90}$$

The ground state energy $e_N(J,h)$ does not depend on N when $c = -1$; approaches monotonically its $\sup_{N\in\mathbb{N}}$ when $c < -1$ or its $\inf_{N\in\mathbb{N}}$ when $c > -1$.

Proof. In the ferromagnetic case the ground state energy is the one that has all the spins aligned with the magnetic field h, that *wlog* can be considered positive.

$$e_N(J,h) = \inf_{\sigma \in \Sigma_N} \frac{H_N(\sigma; J, h)}{N} = -\frac{J}{N(N+c))} \frac{N(N-1)}{2} - h = -\frac{J}{2} \frac{N-1}{N+c} - h \tag{1.91}$$

The claim follows directly from the previous equality. For example, one immediately sees that for $c = -1$, N disappears from the expression of the GS. □

Remark 1.2.1. The fact that pressure and ground state energy are related in the low temperature limit $\beta \to \infty$, and that the latter changes its monotonicity while approaching its own limit for large N lets us think that a similar change in the monotonicity should be observed also in the thermodynamic limit of the pressure per particle, as discussed later.

1.2.2 Normalization effects on finite size corrections

The technique that has been employed to computed the finite size corrections uses a really common strategy that is possible when the interactions are ferromagnetic. We will use a generalized version of it in Chapter 2 to compute finite size corrections for another model.

Theorem 1.2.3 (Normalization effects on finite size corrections)**.** *Consider a system with Hamiltonian:*

$$H_N(\sigma) = \frac{-1}{N+c} \sum_{i,j \in I_N, i<j} \sigma_i \sigma_j \tag{1.92}$$

Then for $\beta > 1$ and sufficiently large N there is a $c^* \in \mathbb{R}$ such that for $c < c^*$, P_N/N converges towards its $\sup_{\mathbb{N}}$, whereas for $c > c^*$ converges towards its $\inf_{\mathbb{N}}$.

Lemma 1.2.4 (Hubbard-Stratonovič transform)**.** *For any $\sigma > 0$ and $t \in \mathbb{R}$,*

$$\exp\left(\frac{t^2\sigma^2}{2}\right) = \int_{\mathbb{R}} \frac{dx}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2} + tx\right) \tag{1.93}$$

Lemma 1.2.5 (Laplace's estimate)**.** *Let $F(z)$ be a twice differentiable function with a global maximum in z^*, and $g(z)$ an analytic function in a neighbourhood of z^*. The following estimate holds:*

$$\int_{\mathbb{R}} dz\, e^{NF(z)} g(z) = \sqrt{\frac{2\pi}{N|F''(z^*)|}} e^{NF(z^*)} \left(g(z^*) + O\left(\frac{1}{N}\right)\right) \tag{1.94}$$

Proof: Theorem 1.2.3. To begin with, we add the symmetric and diagonal terms missing in the Hamiltonian:

$$H_N(\sigma) = -\frac{1}{2(N+c)} \sum_{i,j\in I_N,\, i\neq j} \sigma_i\sigma_j = -\frac{1}{2(N+c)} \sum_{i,j\in I_N} \sigma_i\sigma_j + \frac{N}{2(N+c)} \tag{1.95}$$

We then write it in terms of the configurational magnetization:

$$H_N(\sigma) = -\frac{N^2}{2(N+c)} m_N^2(\sigma) + \frac{N}{2(N+c)} \tag{1.96}$$

With the previous Hamiltonian, the partition function of the system can be written in this form:

$$\begin{aligned} Z_N = \sum_{\tau\in\Sigma_N} \exp\left[\frac{\beta N^2}{2(N+c)} m_N^2(\tau) - \frac{\beta N}{2(N+c)}\right] = \exp\left[-\frac{\beta N}{2(N+c)}\right] \times \\ \times \sum_{\tau\in\Sigma_N} \exp\left[\frac{\beta N^2}{2(N+c)} m_N^2(\tau)\right] \end{aligned} \tag{1.97}$$

In order to obtain a linear term in the magnetization in the exponential, we can use Lemma 1.2.4. If we choose $t = m_N(\tau)$ and $\sigma^2 = \beta N^2/(N+c)$, we get:

$$\begin{aligned} Z_N &= \exp\left[-\frac{\beta N}{2(N+c)}\right] \sum_{\tau\in\Sigma_N} \int_{\mathbb{R}} \frac{dx}{N} \sqrt{\frac{N+c}{2\pi\beta}} \exp\left[-\frac{x^2(N+c)}{2\beta N^2} + m_N(\tau)x\right] = \\ &= \sqrt{\frac{N+c}{2\pi\beta}} \exp\left[-\frac{\beta N}{2(N+c)}\right] \sum_{\tau\in\Sigma_N} \int_{\mathbb{R}} dz \exp\left[-\frac{z^2(N+c)}{2\beta} + Nm_N(\tau)z\right] \end{aligned} \tag{1.98}$$

The dependence on the spin configurations is contained in the exponential of the magnetization. The sum over $\tau \in \Sigma_N$ of those exponential factors is nothing but the partition function of a free system, and z can be seen as the analogous of an inverse absolute temperature. These considerations produce the following result:

$$Z_N = \sqrt{\frac{N+c}{2\pi\beta}} \exp\left[-\frac{\beta N}{2(N+c)}\right] \int_{\mathbb{R}} dz\, e^{NF(z)} g(z) \tag{1.99}$$

$$F(z) = \log(2\cosh z) - \frac{z^2}{2\beta} \tag{1.100}$$

$$g(z) = \exp\left(-\frac{cz^2}{2\beta}\right) \tag{1.101}$$

The integral can be approximated with Laplace's method, illustrated in Lemma 1.2.5. To use this lemma we need to compute the first and second derivatives of F:

$$F'(z) = \tanh z - \frac{z}{\beta} \tag{1.102}$$

$$F''(z) = 1 - \tanh^2 z - \frac{1}{\beta} \tag{1.103}$$

For $\beta < 1$ the unique stationary point is $z^* = 0$, and $F''(0) = 1 - 1/\beta < 0$, thus $F(0)$ is a maximum. Furthermore the F is always concave, so $z^* = 0$ is the global maximum. Since $F(0) = \log 2$, and $g(0) = 1$ we have:

$$\begin{aligned} Z_N &= \sqrt{\frac{N+c}{2\pi\beta}} \exp\left[-\frac{\beta N}{2(N+c)}\right] \sqrt{\frac{2\pi}{N(1/\beta - 1)}} e^{N\log 2} \left(1 + O\left(\frac{1}{N}\right)\right) = \\ &= \sqrt{\frac{N+c}{N(1-\beta)}} \exp\left[-\frac{\beta N}{2(N+c)}\right] 2^N \left(1 + O\left(\frac{1}{N}\right)\right) \end{aligned} \tag{1.104}$$

After taking the logarithm divided by N we have an estimate of the pressure per particle:

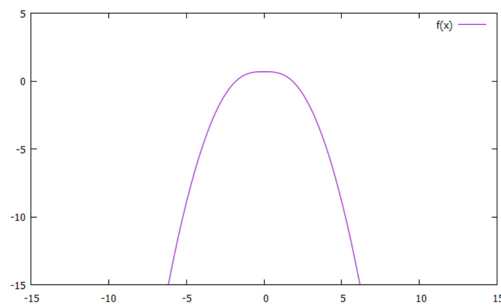
$$\begin{aligned} P_N = \frac{\log Z_N}{N} &= \log 2 - \frac{\beta}{2N}\frac{N}{N+c} - \frac{1}{2N}\left[\log\left((1-\beta)\frac{N}{N+C}\right)\right] + \\ &+ \frac{1}{N}\log\left(1 + O\left(\frac{1}{N}\right)\right) = \log 2 - \frac{1}{N}\frac{1}{2}\left[\beta + \log(1-\beta)\right] + O\left(\frac{1}{N^2}\right) \end{aligned} \tag{1.105}$$

Observe that the coefficient of $1/N$ is always positive thanks to the concavity of the logarithm. This means that, for sufficiently large N, P_N/N approaches its limit from above.

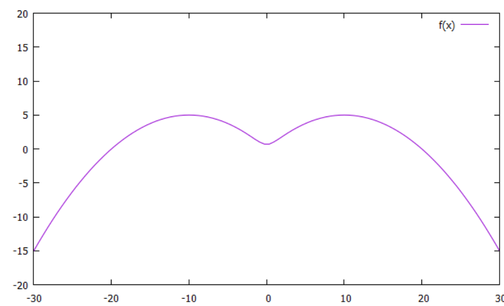
In the case $\beta > 1$ the equation:

$$\tanh z^* = \frac{z^*}{\beta} \tag{1.106}$$

has three solutions. It is immediate to see that $z^* = 0$ is a minimum now, while



(a) $F(z)$ when $\beta < 1$.



(b) $F(z)$ when $\beta > 1$.

there are two new degenerate maxima z^* and $-z^*$, because the function is even, is convex in a neighbourhood of the origin and goes to $-\infty$ as $z \to \pm\infty$.

In case of multiple degenerate maxima, each of them gives a contribution to the integral, that has to be sum with the others. Fortunately the two maxima are symmetric with respect to the origin, and $g(z)$, $F(z)$ and its derivatives are even functions, Calculating F, F'' and g on z^* we get:

$$F(z^*) = \log(2\cosh z^*) - \frac{z^{*2}}{2\beta} \tag{1.107}$$

$$F''(z^*) = 1 - \tanh^2 z^* - \frac{1}{\beta} = 1 - \frac{z^{*2}}{\beta^2} - \frac{1}{\beta} \tag{1.108}$$

$$g(z^*) = \exp\left(-\frac{cz^{*2}}{2\beta}\right) \tag{1.109}$$

Finally, inserting these values in our estimate of Z_N:

$$Z_N = 2\sqrt{\frac{N+c}{N(1-\beta+z^{*2}/\beta)}}\exp\left[-\frac{\beta N}{2(N+c)} + N\left(\log(2\cosh z^*) - \frac{z^{*2}}{2\beta}\right)\right] \times \\ \times \left[\exp\left(-\frac{cz^{*2}}{2\beta}\right) + O\left(\frac{1}{N}\right)\right] \tag{1.110}$$

We calculate again the logarithm divided by N and segregating the orders in $1/N$:

$$P_N = \frac{\log Z_N}{N} = \log(2\cosh z^*) - \frac{z^{*2}}{2\beta} + \frac{K}{N} + O\left(\frac{1}{N^2}\right) \tag{1.111}$$

$$K = \log 2 - \frac{1}{2}\left[\beta + \log(1-\beta+z^{*2}/\beta) + \frac{cz^{*2}}{\beta}\right] \tag{1.112}$$

Contrary to what we have proved before, here the coefficient of the $1/N$ order has not a definite sign. Its sign depends on the value we choose for c. For $c > 0$ sufficiently large, say $c > c^*$, there is a different asymptotic behaviour: P_N/N converges to its $\inf_{\mathbb{N}}$ from above.

□

1.3 Solution of the ferromagnetic CW model

Here we propose the solution of the CW ferromagnetic model as carried out by Francesco Guerra in [16], Theorem 1.

Theorem 1.3.1 (TDL of CW)**.** *The limit as $N \to \infty$ of:*

$$p_N(\beta; J, h) = \frac{1}{N}\log\sum_{\sigma\in\Sigma_N}\exp\left[\beta N\left(\frac{J}{2}m_N^2(\sigma) + hm_N(\sigma)\right)\right] \tag{1.113}$$

is:

$$p(\beta; J, h) = \sup_{x\in\mathbb{R}}\left[-\frac{\beta J x^2}{2} + \log 2 + \log\cosh(\beta(Jx+h))\right] \tag{1.114}$$

Proof. We use the fact that $(M - m_N)^2 \geq 0 \;\Rightarrow\; m_N^2 \geq 2Mm_N - M^2$.

$$\begin{aligned} p_N(\beta; J, h) &\geq \frac{1}{N}\log\sum_{\sigma\in\Sigma_N}\exp\left[\beta N\left(-\frac{JM^2}{2} + JMm_N(\sigma) + hm_N(\sigma)\right)\right] = \\ &= -\frac{\beta JM^2}{2} + \frac{1}{N}\log\sum_{\sigma\in\Sigma_N}\exp\left[\beta N m_N(\sigma)\left(JM+h\right)\right] \end{aligned} \tag{1.115}$$

The argument of the exponential is linear in the magnetization, this signals a free system. We use the solution we already know:

$$p_N(\beta; J, h) \geq -\frac{\beta JM^2}{2} + \log 2 + \log\cosh[\beta(JM+h)] \tag{1.116}$$

We take the $\sup_{M\in\mathbb{R}}$ of the r.h.s.

$$p_N(\beta; J, h) \geq \sup_{M\in\mathbb{R}} \left[-\frac{\beta J M^2}{2} + \log 2 + \log\cosh(\beta(JM + h)) \right] \tag{1.117}$$

For the upper bound we take $M \in S_N$, the spectrum of the magnetization, which has $N+1$ elements and we introduce the sum over a Kronecker delta.

$$\begin{aligned} Z_N(\beta; J, h) &= \sum_{\sigma\in\Sigma_N} \sum_{M\in S_N} \delta_{Mm_N} \exp\left[\beta N \left(\frac{J}{2} m_N^2(\sigma) + h m_N(\sigma)\right)\right] = \\ &= \sum_{M\in S_N} \sum_{\sigma\,:\,m_N(\sigma)=M} \exp\left[\beta N \left(-\frac{JM^2}{2} + JM m_N(\sigma) + h m_N(\sigma)\right)\right] \end{aligned} \tag{1.118}$$

Now we shall take the $\sup_{M\in\mathbb{R}}$ of the exponent and extend the sum to the whole configuration space Σ_N.

$$\begin{aligned} Z_N &\leq \sum_{M\in S_N} \exp \sup_{x\in\mathbb{R}} \left[N\log 2 + N\log\cosh(\beta(Jx + h)) - N\frac{\beta J x^2}{2} \right] = \\ &= (N+1) \exp \sup_{x\in\mathbb{R}} \left[N\log 2 + N\log\cosh(\beta(Jx + h)) - N\frac{\beta J x^2}{2} \right] \end{aligned} \tag{1.119}$$

After taking the logarithm of the expression we have also an upper bound to the pressure per particle. We set:

$$p(x; \beta; J, h) = -\frac{\beta J x^2}{2} + \log 2 + \log\cosh(\beta(Jx + h)) \tag{1.120}$$

To sum up:

$$\sup_{x\in\mathbb{R}} p(x; \beta; J, h) \leq p_N(\beta; J, h) \leq \frac{\log(N+1)}{N} + \sup_{x\in\mathbb{R}} p(x; \beta; J, h) \tag{1.121}$$

hence:

$$\lim_{N\to\infty} p_N(\beta; J, h) = \sup_{x\in\mathbb{R}} p(x; \beta; J, h) = p(\bar{x}; \beta; J, h) \tag{1.122}$$

$\square$

Proposition 1.3.2 (Meaning of $\bar{x}$). *$\bar{x}$ is the magnetization of the system in the thermodynamic limit.*

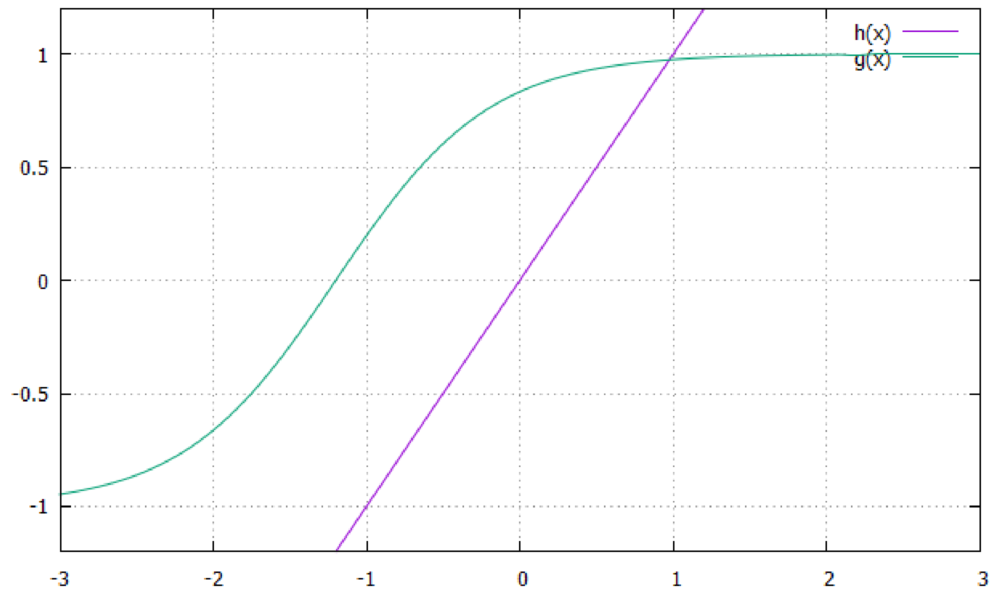


Figure 1.1: Graphic solution of (1.123) for $J = 1$ and $h = 1.2$.

Proof. Having absorbed β in the parameters the equation for $\bar{x}$ reads:

$$\frac{\partial p}{\partial x}(\bar{x}; J, h) = -J\bar{x} + \tanh(J\bar{x} + h)J = 0 \quad \Rightarrow \quad \bar{x} = \tanh(J\bar{x} + h) \tag{1.123}$$

Now we compute the first derivative of the pressure in the thermodynamic limit, assuming to be able to exchange the derivative with the limit:

$$\begin{aligned}\lim_{N\to\infty} \frac{\partial}{\partial h} p_N(J, h) = \frac{\partial}{\partial h} p(\bar{x}(J, h); J, h) = -J\bar{x}\frac{\partial \bar{x}}{\partial h} + \tanh(J\bar{x} + h)\left[1 + J\frac{\partial \bar{x}}{\partial h}\right] = \\ = -J\bar{x}\frac{\partial \bar{x}}{\partial h} + \bar{x} + \bar{x}J\frac{\partial \bar{x}}{\partial h} = \bar{x}\end{aligned} \tag{1.124}$$

A graphic representation of the solution is represented in Figure1.1. □

1.4 Antiferromagnetic Curie-Weiss model

Proposition 1.4.1 (Extensivity oh the Hamiltonian)**.** *Consider the Hamiltonians* (1.9) *and* (1.52) *with* $J < 0$*. There are* $\underline{K}, \bar{K} \in \mathbb{R}$ *such that* $\underline{K}N \leq H_N \leq \bar{K}N$*.*

Proof. Let us consider (1.9) for the moment.

$$\begin{aligned} H_N^{CW}(\sigma) = \frac{-J}{N-1} \sum_{i,j\in I_N, i<j} \sigma_i\sigma_j - h\sum_{i\in I_N}\sigma_i = \\ = \frac{-J}{2(N-1)} \sum_{i,j\in I_N} \sigma_i\sigma_j - h\sum_{i\in I_N}\sigma_i + \frac{NJ}{2(N-1)} = \\ = \frac{NJ}{2(N-1)} - \frac{JN^2}{2(N-1)} m_N^2(\sigma) - hNm_N(\sigma) \leq \\ \leq -\frac{NJ(N-1)}{2(N-1)} + |h|N = N \underbrace{\left(\frac{|J|}{2} + |h|\right)}_{\bar{K}} \end{aligned} \tag{1.125}$$

Analogously:

$$H_N^{CW}(\sigma) \geq \frac{NJ}{2(N-1)} - |h|N = -N\left(\frac{|J|}{2(N-1)} + |h|\right) \geq N \underbrace{\left(-\frac{|J|}{2} - |h|\right)}_{\underline{K}} \tag{1.126}$$

Now we turn to (1.52).

$$H_N^{CW}(\sigma) = -JNm_N^2(\sigma) - Nhm_N(\sigma) \leq N \underbrace{(|J| + |h|)}_{\bar{K}} \tag{1.127}$$

$$H_N^{CW}(\sigma) \geq N(-|h|) = N\underline{K} \tag{1.128}$$

□

Corollary 1.4.2. *Te pressure per particle of* (1.9) *and* (1.52) *are bounded from above and below.*

Proof. It is a consequence of the previous proposition and of Proposition 1.1.3. □

Remark 1.4.1. Notice that once the bounds for the pressure are given, it is still possible to prove sub- or super-additivity with the tools developed in the previous sections. For example, the pressure generated by (1.9) is always sub-additive thanks to the symmetrization lemma, in which the sign of J is irrelevant.

On the contrary, the sign of J determines whether the pressure is super- or sub-additive when using the Guerra-Toninelli interpolation. In the anti-ferromagnetic case, with Hamiltonian (1.52) the resulting pressure turns out to be super-additive.

To summarize, while in the case (1.9) the pressure per particle approaches its $\inf_{\mathbb{N}}$ when $N \to \infty$, with (1.52) the pressure per particle approaches its $\sup_{\mathbb{N}}$.

However the two pressures in the thermodynamic limit must coincide, because the initial Hamiltonians differ for terms that are thermodynamically irrelevant.

Theorem 1.4.3 (Pressure per particle, $N-1$ normalization, $h=0$)**.** *Consider the Hamiltonian:*

$$H_N^{CW} = \frac{-J}{N-1} \sum_{i,j \in I_N,\, i<j} \sigma_i \sigma_j \quad \textit{with} \quad J < 0 \tag{1.129}$$

The pressure per particle of this model has a trivial limit:

$$p_N(\beta; J) = \frac{P_N(\beta; J)}{N} \longrightarrow \log 2 \quad \textit{when } N \to \infty \tag{1.130}$$

Proof. Let us write the pressure per particle explicitly:

$$p_N = \frac{1}{N} \log \sum_{\sigma \in \Sigma_N} \exp\left(\frac{-\beta |J|}{N-1} \sum_{i,j \in I_N,\, i<j} \sigma_i \sigma_j \right) \tag{1.131}$$

Now we use the elementary inequality:

$$0 \leq \sum_{i,j \in I_N} \sigma_i \sigma_j = N^2 m_N^2(\sigma) = N + 2 \sum_{i<j} \sigma_i \sigma_j \quad \Rightarrow \quad \sum_{i<j} \sigma_i \sigma_j \geq -\frac{N}{2} \tag{1.132}$$

This brings us to:

$$p_N \leq \frac{1}{N} \log \sum_{\sigma \in \Sigma_N} \exp\left(\frac{N\beta |J|}{2(N-1)} \right) = \frac{1}{N} \log \left[2^N \exp\left(\frac{N\beta|J|}{2(N-1)} \right) \right] =$$
$$= \log 2 + \frac{\beta |J|}{2(N-1)} \tag{1.133}$$

Once the upper bound is found, we turn to the lower bound. In order to find it we notice that the pressure can always be written in this form:

$$p_N = \frac{1}{N} \log 2^N \omega_0 \left[\exp\left(\frac{-\beta |J|}{N-1} \sum_{i,j \in I_N,\, i<j} \sigma_i \sigma_j \right) \right] \quad \text{with} \quad \omega_0(\,\cdot\,) = \frac{1}{2^N} \sum_{\sigma \in \Sigma_N} (\,\cdot\,) \tag{1.134}$$

ω_0 is the uniform probability measure over the configuration space. Now, since the exponential is a convex function we can use Jensen's inequality:

$$p_N \geq \frac{1}{N} \log 2^N \exp\left(\frac{-\beta|J|}{N-1} \sum_{i,j\in I_N,\, i<j} \omega_0(\sigma_i\sigma_j)\right) = \log 2 \tag{1.135}$$

In the last step we have used the fact that the uniform measure factorizes and that each spin has a vanishing expectation with respect to the latter.

Finally:

$$\log 2 \leq p_N(\beta; J) \leq \log 2 + \frac{\beta|J|}{2(N-1)} \tag{1.136}$$

Hence $p_N \longrightarrow \log 2$ as $N \to \infty$. □

Remark 1.4.2. Notice that the pressure goes towards its $\inf_{\mathbb{N}}$, as predicted by the symmetrization lemma. We expect that, in the following second case, the pressure will approach its $\sup_{\mathbb{N}}$ in the thermodynamic limit.

Theorem 1.4.4 (Pressure per particle, N normalization, $h=0$)**.** *Consider the Hamiltonian:*

$$H_N^{CW} = \frac{-J}{N} \sum_{i,j\in I_N} \sigma_i\sigma_j = |J|Nm_N^2(\sigma) \quad \textit{with} \quad J<0 \tag{1.137}$$

The pressure per particle of this model has a trivial limit:

$$p_N(\beta; J) = \frac{P_N(\beta; J)}{N} \longrightarrow \log 2 \quad \textit{when } N \to \infty \tag{1.138}$$

Proof. The path is similar to that of the previous proof. Let us start with the lower bound.

$$\begin{aligned} p_N &= \frac{1}{N} \log \sum_{\sigma\in\Sigma_N} \exp\left(-\beta|J|Nm_N^2(\sigma)\right) = \frac{1}{N} \log 2^N \omega_0 \left[\exp\left(\frac{-\beta|J|}{N} \sum_{i,j\in I_N} \sigma_i\sigma_j\right)\right] \\ &\geq \frac{1}{N} \log 2^N \exp\left(\frac{-\beta|J|}{N} \sum_{i,j\in I_N} \underbrace{\omega_0(\sigma_i\sigma_j)}_{\delta_{ij}}\right) = \frac{1}{N} \log 2^N \exp\left(-\beta|J|\right) = \log 2 - \frac{\beta|J|}{N} \end{aligned} \tag{1.139}$$

The upper bound is much easier:

$$p_N = \frac{1}{N} \log \sum_{\sigma\in\Sigma_N} \exp\left(-\beta|J|Nm_N^2(\sigma)\right) \leq p_N = \frac{1}{N} \log \sum_{\sigma\in\Sigma_N} 1 = \log 2 \tag{1.140}$$

Resuming the results we have:

$$\log 2 - \frac{\beta|J|}{N} \leq p_N(\beta; J) \leq \log 2 \tag{1.141}$$

hence $p_N \longrightarrow \log 2$ as $N \to \infty$. □

There are now two ways to proceed in the analysis of the antiferromagnetic model with $h \neq 0$. The following one, presented by A. Bovier in [9], is also valid in the ferromagnetic case and makes use of Stirling approximation. In the last section of this chapter we will show an alternative way to find bounds for the pressure of the antiferromagnetic model. The technique developed will be used also for multi-species and multi-layer systems in Chapter 3.

1.4.1 Solution via Stirling approximation

Remark 1.4.3. Thanks to the fact that, at least in the case (1.52), the Hamiltonian depends on the spin only through the magnetization, the partition function can also be computed as a sum over the possible values of the magnetization $m \in S_N = \{-1, -1+2/N, ..., 1-2/N, +1\}$.

$$Z_N = \sum_{\sigma\in\Sigma_N} \exp\left(\beta N\psi(m_N(\sigma); J, h)\right) = \sum_{m\in S_N} z_{m,N} \exp\left(\beta N\psi(m; J, h)\right) \tag{1.142}$$

where $z_{m,N}$ is a binomial coefficient. In order to find it one has to count the possible configurations that produce a magnetization equal to m. If N_+ are the spins with value $+1$ and N_- the ones with value -1 then:

$$\begin{cases} Nm = N_+ - N_- \\ N = N_+ + N_- \end{cases} \Rightarrow \begin{cases} N_+ = \frac{N(1+m)}{2} \\ N_- = \frac{N(1-m)}{2} \end{cases} \tag{1.143}$$

The binomial coefficient is:

$$z_{m,N} = \frac{N!}{\left(\frac{N(1+m)}{2}\right)! \left(\frac{N(1-m)}{2}\right)!} \tag{1.144}$$

Lemma 1.4.5 (Asymptotic behaviour of $z_{m,N}$)**.**

$$\frac{1}{N}\log z_{m,N} = \log 2 - I(m) - J_N(m) \tag{1.145}$$

where:

$$I(m) = \frac{1}{2}\left[(1+m)\log(1+m) + (1-m)\log(1-m)\right] \tag{1.146}$$

$$J_N(m) = \frac{1}{2N}\log\frac{\pi N(1-m^2)}{2} + O\left(\frac{1}{N^2}\left(\frac{1}{1+m}+\frac{1}{1-m}\right)\right) \tag{1.147}$$

Proof. Thanks to Stirling's formula we have:

$$\begin{aligned} z_{m,N} = \frac{\sqrt{2\pi N}(N/e)^N\left[1+O\left(\frac{1}{N}\left(\frac{1}{1+m}+\frac{1}{1-m}\right)\right)\right]}{\sqrt{\pi N(1-m)\pi N(1+m)}\left(\frac{N(1+m)}{2e}\right)^{N(1+m)/2}\left(\frac{N(1-m)}{2e}\right)^{N(1-m)/2}} = \\ = \sqrt{\frac{2}{\pi N(1-m^2)}}2^N(1+m)^{-N(1+m)/2}(1-m)^{-N(1-m)/2}\times \\ \times\left[1+O\left(\frac{1}{N}\left(\frac{1}{1+m}+\frac{1}{1-m}\right)\right)\right] \end{aligned} \tag{1.148}$$

Taking the logarithm divided by N of the previous formula yields:

$$\begin{aligned} \frac{\log z_{m,N}}{N} = \log 2 - \frac{1}{2}\left[(1+m)\log(1+m)+(1-m)\log(1-m)\right] - \\ - \frac{1}{2N}\log\frac{\pi N(1-m^2)}{2} + O\left(\frac{1}{N^2}\left(\frac{1}{1+m}+\frac{1}{1-m}\right)\right) \end{aligned} \tag{1.149}$$

□

Theorem 1.4.6 (Pressure per particle with $h \neq 0$, N normalization)**.** *Consider the Hamiltonian* (1.52)*. Its pressure in the thermodynamic limit is:*

$$p(\beta;J;h) = \lim_{N\to\infty}\frac{\log Z_N}{N} = \sup_{m\in[-1,1]} A(m;\beta;J,h) \tag{1.150}$$

where:

$$A(m;\beta;J,h) = \frac{-\beta|J|m^2}{2} + \beta hm + \log 2 - I(m) \tag{1.151}$$

Sketch of the proof: We rewrite the partition function thanks to the previous lemma in the following way:

$$Z_N = \sum_{m \in S_N} \exp\left[N\left(\frac{-\beta|J|m^2}{2} + \beta hm + \log 2 - I(m) - J_N(m)\right)\right] \tag{1.152}$$

In this sum there is a term that dominates the others when N becomes large, and this very term will be the only significant contribution to the partition function, and corresponds to the magnetization that maximizes the exponential:

$$\frac{\log Z_N}{N} = \max_{m \in S_N}\left(\frac{-\beta|J|m^2}{2} + \beta hm + \log 2 - I(m)\right) + O(\log N/N) \tag{1.153}$$

where $O(\log N/N)$ comes from $J_N(m)$. The set on which we seek for the maximum depends on N, but as shown in [9], we can alternatively compute the $\sup_{m\in[-1,1]}$ at the cost of adding another $O(\log N/N)$. More precisely:

$$\begin{aligned}\max_{m \in S_N}\Bigg|\left(\frac{-\beta|J|m^2}{2} + \beta hm + \log 2 - I(m)\right) - \\ - \sup_{m' \in[-1,1]:|m'-m|\le 2/N}\left(\frac{-\beta|J|m^2}{2} + \beta hm + \log 2 - I(m)\right)\Bigg| \le C\frac{\log N}{N}\end{aligned} \tag{1.154}$$

Finally we are free to perform the limit $N \to \infty$:

$$\begin{aligned}p_N(\beta; J, h) = \sup_{m\in[-1,1]}\left(\frac{-\beta|J|m^2}{2} + \beta hm + \log 2 - I(m)\right) + O(\log N/N) \longrightarrow \\ \longrightarrow \sup_{m\in[-1,1]}\left(\frac{-\beta|J|m^2}{2} + \beta hm + \log 2 - I(m)\right) = \sup_{m\in[-1,1]} A(m; \beta; J, h)\end{aligned} \tag{1.155}$$

□

Remark 1.4.4. The terms in the expression of (1.151) have an interesting interpretation. Let us start with $-I(m)$. It comes from the binomial coefficients and basically counts the number of configurations compatible with a given magnetization m, so we can say that is an entropic term. Since $-I(m)$ is concave and has a maximum in $m = 0$, it favours little magnetizations. The same can be said for $-\beta|J|m^2/2$. The only term that shifts the maximum form $m = 0$ is the magnetic field term βhm. A plot of A for certain values β, J, h is shown in Figure1.2.

Moreover, for high temperatures $\beta \to 0$ the entropic terms is the only relevant and, as it should be, the maximum approaches $m = 0$.

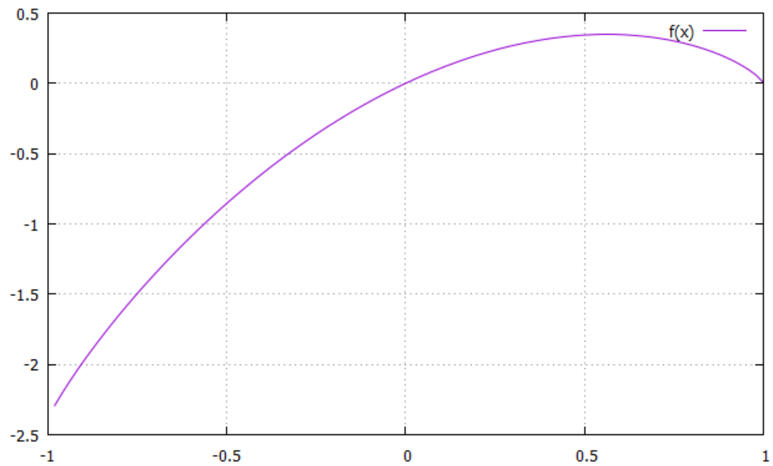


Figure 1.2: $A(m;\beta,J,h)$ for $\beta|J|=1$ and $\beta h=1.2$.

Corollary 1.4.7. *The* $\bar{m}$ *that realizes the* $\sup_{m\in[-1,1]}$ *in* (1.150) *satisfies:*

$$-\bar{m}=\tanh\beta(|J|\bar{m}-h) \tag{1.156}$$

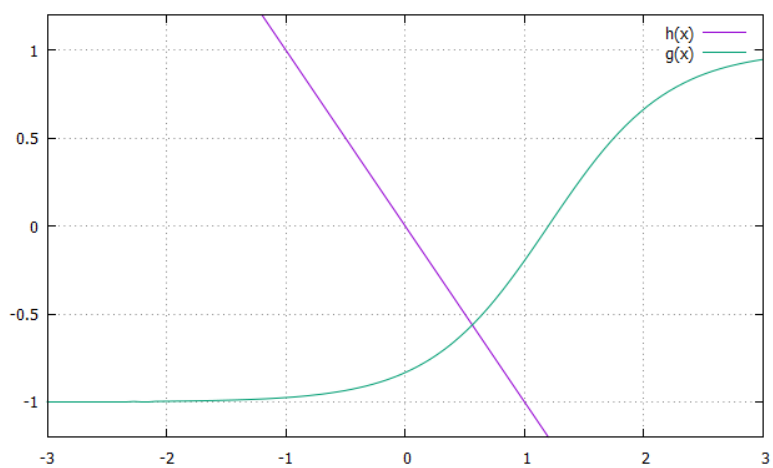


Figure 1.3: Graphic solution of (1.156) for $\beta|J|=1$ and $\beta h=1.2$.

Proof. Computing the first derivative of (1.150), which is concave, we get:

$$\frac{dA}{dm}(\bar{m})=-\tanh^{-1}\bar{m}-\beta|J|\bar{m}+\beta h=0 \quad\Rightarrow\quad \bar{m}=\tanh\beta(h-|J|\bar{m}) \tag{1.157}$$

$\square$

1.4.2 Solution via interpolation

There is a more elegant way to obtain the limit of the CW antiferromagnetic model with a non vanishing magnetic external field h. Consider the following lemma:

Lemma 1.4.8. *Let $p_N(t)$ be the interpolating generating functional, induced by:*

$$H_N(t) = tH_N + (1-t)\tilde{H}_N(x) \tag{1.158}$$

where x is a collection of parameters, namely:

$$p_N(t) = \frac{1}{N}\log\sum_{\sigma\in\Sigma}\exp(-H_N(t)) \tag{1.159}$$

Then $p_N(t)$ is convex in t, in particular:

$$p'_N(0) \leq p'_N(t) \leq p'_N(1) \tag{1.160}$$

Proof. The proof basically consists in calculating the second derivative of $p_N(t)$.

$$p'_N(t) = -\frac{1}{N}\omega_{N,t}\left(H_N - \tilde{H}_N\right) \tag{1.161}$$

$$p''_N(t) = \frac{1}{N}\omega_{N,t}\left[\left(H_N - \tilde{H}_N\right)^2\right] - \frac{1}{N}\omega^2_{N,t}\left(H_N - \tilde{H}_N\right) \geq 0 \tag{1.162}$$

The positivity of the second derivative follows from the positivity of the variance. □

Theorem 1.4.9 (Solution of the antiferromagnetic CW model)**.** *Given the hamiltonian:*

$$H^{af}_N = -\frac{J}{2}Nm^2_N - Nhm_N \quad \text{with } J < 0 \tag{1.163}$$

then the pressure per particle in the thermodynamic limit is:

$$p^{CW}_{af} = \lim_{N\to\infty}\frac{1}{N}\log\sum_{\sigma\in\Sigma_N}\exp(-H^{af}_N) = \inf_{x\in\mathbb{R}}\left[-\frac{J}{2}x^2 + \log 2\cosh\left(Jx+h\right)\right] \tag{1.164}$$

Proof. Let us start by defining:

$$\tilde{H}_N = -Nm_N(Jx+h) \tag{1.165}$$

$$H_N(t) = tH^{af}_N + (1-t)\tilde{H}_N \tag{1.166}$$

$$p_N(t) = \frac{1}{N}\log\sum_{\sigma\in\Sigma_N}\exp(-H_N(t)) \tag{1.167}$$

By doing so we can explicitly compute the interpolating pressure at the instant $t = 0$. Moreover, at the instant $t = 1$ we have the finite size pressure.

$$p_N(0) = p_N^{Free} = \log 2\cosh(Jx + h) \tag{1.168}$$
$$p_N(1) = p_{N,af}^{CW} \tag{1.169}$$

Applying the fundamental theorem of the integral calculus we get:

$$p_{N,af}^{CW} = \log 2\cosh(Jx + h) + \int_0^1 p_N'(t)\,dt \tag{1.170}$$
$$p_N'(t) = \omega_{N.t}\left(-\frac{1}{N}\frac{dH(t)}{dt}\right) \tag{1.171}$$

Let us focus on the derivative of the interpolating hamiltonian:

$$-\frac{1}{N}\frac{dH(t)}{dt} = \omega_{N,t}\left(\frac{J}{2}m_N^2 - Jxm_N\right) = \frac{J}{2}\omega_{N,t}[(m_N - x)^2] - \frac{Jx^2}{2} \tag{1.172}$$

So we are left with an expression that immediately gives us an upper bound, thanks to the fact that we have completed a square multiplied by J, which is negative.

$$\begin{aligned} p_{N,af}^{CW} &= \log 2\cosh(Jx + h) - \frac{Jx^2}{2} + \frac{J}{2}\int_0^1 dt\,\omega_{N,t}\left[(m_N - x)^2\right] \le \\ &\le \log 2\cosh(Jx + h) - \frac{Jx^2}{2} = A(x; J, h) \end{aligned} \tag{1.173}$$

For all values of J and h, $A(x; J, h)$ is convex in x, thus we can safely optimize the bound by taking the $\inf_x$:

$$p_{N,af}^{CW} \le \inf_x A(x; J, h) = A(\bar{x}; J, h) \tag{1.174}$$
$$\bar{x} = \tanh(J\bar{x} + h) \tag{1.175}$$

The lower bound can be found thanks to the previous lemma.

$$p_{N,af}^{CW} \ge p_N(0) + \int_0^1 dt\,p_N'(0) = \log 2\cosh(Jx + h) - \frac{Jx^2}{2} + \frac{J}{2}\omega_{N,0}\left[(m_N - x)^2\right] \tag{1.176}$$

The expectation in the product measure $\omega_{N,0}$, which is related to the free system, can be explicitly calculated starting from this elementary result:

$$\omega_{N,0}(m_N) = \omega_{N,0}(\sigma_1) = \tanh(Jx + h) \tag{1.177}$$

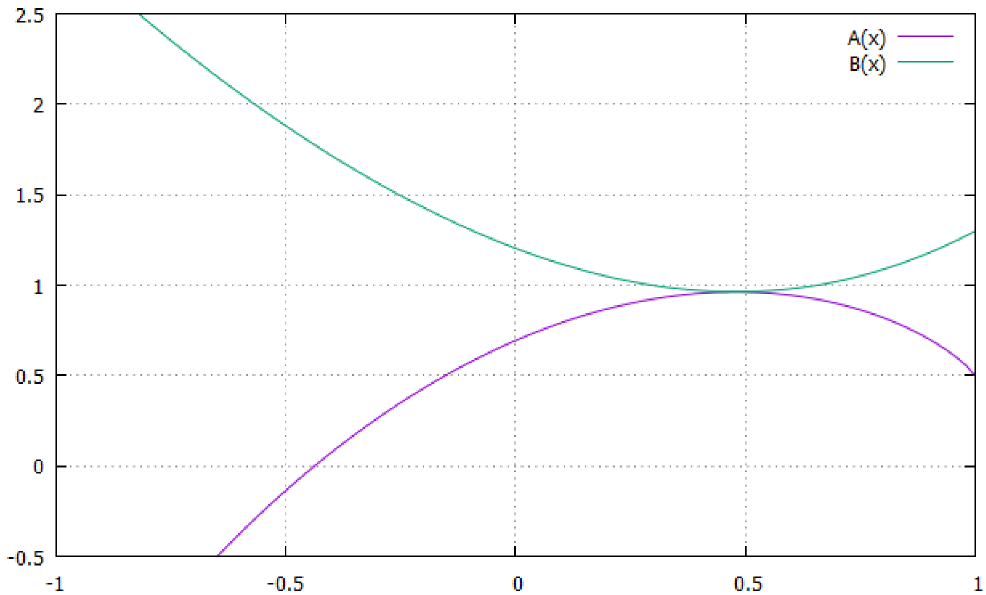


Figure 1.4: The two variational functions compared.

Once the diagonal terms in the square magnetization m_N^2 are segregated one can proceed:

$$\omega_{N,0}\left(m_N^2 - 2xm_N + x^2\right) = \omega_{N,0}\left(\frac{1}{N} + \frac{1}{N^2}\sum_{i\neq j}\sigma_i\sigma_j - 2x\frac{1}{N}\sum_i \sigma_i + x^2\right) = $$
$$= \frac{1}{N} + \frac{N-1}{N}\tanh^2(Jx+h) - 2x\tanh(Jx+h) + x^2 \quad (1.178)$$

With the specific choice $x = \bar{x} = \tanh(J\bar{x}+h)$ the previous expression simplifies to:

$$\omega_{N,0}\left[(m_N - \bar{x})^2\right] = \frac{1-\bar{x}^2}{N} \quad (1.179)$$

We can finally conclude that:

$$p_{N,af}^{CW} \geq \log 2\cosh(J\bar{x}+h) - \frac{J\bar{x}^2}{2} + \frac{J(1-\bar{x}^2)}{2N} = \inf_x A(x;J,h) + \frac{J(1-\bar{x}^2)}{2N} \quad (1.180)$$

The lower bound clearly reduces to the upper bound in the thermodynamic limit and this proves the claim. □

Remark 1.4.5. It may seem that the two results we have obtained so far for the antiferromagnetic CW model are in contrast with each other. Fortunately they are not, in fact the two solutions are linked by a Legendre transform of the entropic term $I(m)$. This also explains why in one case we have a sup of a variational expression and an inf in the other case. The two variational expressions calculated in their extremal points are thus equal, as shown in Figure1.4.

Chapter 2

The Sherrington-Kirkpatrick model

In this chapter we will list and give some proofs of a series of important results concerning the SK model (see [23], [18]) which is a disordered version of the CW mean field model treated before, *i.e.* with random interactions *i.i.d.* sampled from a standard gaussian $\mathcal{N}(0,1)$. As widely known, Sherrington and Kirkpartick found the so called *replica symmetric* solution that did not hold for any value of the inverse temperature. In particular, the entropy computed from that solution is negative in the low-temperature limit.

In a series of articles (see [20], [21], [22]), G. Parisi found a *replica symmetry breaking* solution, named in this way for historical reasons, that we will discuss in detail later, introducing a collection of infinite order parameters. Although it was immediately accepted by physicists, for it was in extremely good agreement with simulations, a formal proof was achieved by Guerra [14] (2002), that found the upper bound, and Talagrand [26](2006), that completed the proof with the lower bound. The proof was further simplified by Panchenko in his monograph [18]. In order to prove both upper and lower bounds one needs to represent the Parisi functional for the free energy in terms of *Ruelle Probability Cascades* (RPC). In this chapter we will illustrate only the complete proof for the upper bound, and the main steps that led to the lower bound, as done by D. Panchenko, in addition to the proof of the existence of the limit ([17]).

2.1 The model

Definition 2.1.1 (SK Hamiltonian)**.** Consider a system in which interactions between couple of spins are random variables, more precisely $J_{ij} \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$. The mean field spin glass SK model is defined by the hamiltonian:

$$H_N^{SK}(\sigma; J, h) = -K_N(\sigma; J) - h\sum_{i=1}^{N}\sigma_i \tag{2.1}$$

$$K_N(\sigma; J) = \frac{1}{\sqrt{N}}\sum_{i,j=1}^{N} J_{ij}\sigma_i\sigma_j \tag{2.2}$$

Similarly to the CW model, in the SK model there is a quantity that plays the role of magnetization, called overlap defined as follows.

Definition 2.1.2. Let $\sigma, \tau \in \Sigma_N$ be two configurations of spins. Their *overlap* is:

$$q_N(\sigma, \tau) = \frac{1}{N}\sum_{i=1}^{N}\sigma_i\tau_i \tag{2.3}$$

Let us start with a lemma that will be the key to various proofs.

Lemma 2.1.1 (Integration by parts in gaussian processes)**.** *Consider a centered gaussian family $(X_i)_{i\leq n}$ and a function $F \in C^1(\mathbb{R}^n)$ such that:*

$$\lim_{|x|\to\infty} |F(X)|e^{-a|x|^2} = 0 \quad \forall\, a > 0 \tag{2.4}$$

then:

$$\mathbb{E}_X[X_iF(X)] = \sum_{j=1}^{n} C_{ij}\mathbb{E}_X\left[\frac{\partial F}{\partial X_j}(X)\right] \tag{2.5}$$

where $C_{ij} = \mathbb{E}_X[X_iX_j]$.

Remark 2.1.1. If we treat the configuration of spins as a label, like i or j in the previous lemma, the set $(K_\sigma)_{\sigma\in\Sigma_N}$ becomes a gaussian family, thus it will be completely described by its covariance matrix:

$$C_{\sigma\tau} = \mathbb{E}[K_\sigma K_\tau] = \frac{1}{N}\sum_{i,j,k,l=1}^{N}\mathbb{E}[J_{ij}J_{kl}]\sigma_i\sigma_j\tau_k\tau_l = \frac{1}{N}\sum_{i,j=1}^{N}\sigma_i\sigma_j\tau_i\tau_j = Nq_N^2(\sigma,\tau) \tag{2.6}$$

The disorder introduced in the Hamiltonian makes the pressure itself a random quantity. However, thanks to the well known self-averaging property of the latter, we can consider its quenched version, namely:

$$P_N^{SK} = \mathbb{E} \log \left(\sum_{\sigma \in \Sigma_N} e^{-\beta H_N^{SK}(\sigma)} \right) \tag{2.7}$$

where $\mathbb{E}$ denotes the expectation over the disorder. In the thermodynamic limit the details of the disorder become irrelevant and the pressure per particle will concentrate, as a r.v., around its expectation (see [10]).

2.1.1 Existence of the thermodynamic limit

Theorem 2.1.2 (Existence of the thermodynamic limit, SK)**.** *The quenched pressure of the SK model is super-additive: $P_{N+M} \geq P_N + P_M$, hence the limit of $P_N^{SK}/N = p_N^{SK}$ for $N \to \infty$ exists and:*

$$\lim_{N \to \infty} p_N^{SK} = \sup_N p_N^{SK} \tag{2.8}$$

Proof. The result is achieved through interpolation. Consider a bipartition of the system $\Lambda = \Lambda_1 \uplus \Lambda_2$ with $|\Lambda| = N = |\Lambda_1| + |\Lambda_2| = N_1 + N_2$, where Λ is simply $\{1, 2, \ldots, N\}$. Define:

$$-H_N(\sigma; t) = \underbrace{\sqrt{t} K_\Lambda(\sigma) + \sqrt{1-t}[K_{\Lambda_1}(\sigma) + K_{\Lambda_2}(\sigma)]}_{K_\sigma(t)} + h \sum_{i \in \Lambda} \sigma_i \tag{2.9}$$

$$K_{\Lambda_s}(\sigma) = \frac{1}{\sqrt{2N}} \sum_{i,j \in \Lambda_s} \tilde{J}_{ij} \sigma_i \sigma_j \quad s = 1, 2 \tag{2.10}$$

$$W_\Lambda(\sigma; \beta, h) = e^{\beta h \sum_{i \in \Lambda} \sigma_i} \tag{2.11}$$

$$P_N(t) = \mathbb{E} \log \left(\sum_{\sigma \in \Sigma_N} W_\Lambda(\sigma; \beta, h) e^{\beta K_\sigma(t)} \right) = \mathbb{E} \log Z_N(t) \tag{2.12}$$

where $\tilde{J}$ are independent gaussian processes, and the expectation in the last line is taken with respect to all the r.v. involved.

Now we proceed in the calculation of the first derivative of $P_N(t)$. Let us denote for the moment: $C^{(s)}_{\sigma\tau} = \mathbb{E}[K_{\Lambda_s}(\sigma)K_{\Lambda_s}(\tau)] = N_s q^2_{N,s}(\sigma,\tau)$. then

$$P'_N(t) = \beta\mathbb{E}\left[\frac{\sum_{\sigma\in\Sigma_N} W_\Lambda(\sigma;\beta,h)e^{\beta K_\sigma(t)}}{Z_N(t)}\left(\frac{K_\Lambda(\sigma)}{2\sqrt{t}} - \frac{K_{\Lambda_1}(\sigma)+K_{\Lambda_2}(\sigma)}{2\sqrt{1-t}}\right)\right] =$$

$$= \frac{\beta^2}{2}\mathbb{E}\Omega^{(2)}_{N,t}\left[C_{\sigma\sigma} - C_{\sigma\tau} - C^{(1)\sigma\sigma} + C^{(1)}_{\sigma\tau} - C^{(2)}_{\sigma\sigma} + C^{(2)}_{\sigma\tau}\right] =$$

$$= \frac{\beta^2}{2}\mathbb{E}\Omega^{(2)}_{N,t}\left[N - Nq^2_N(\sigma,\tau) - N_1 + N_1q^2_{N,1}(\sigma,\tau) - N_2 + N_2q^2_{N,2}(\sigma,\tau)\right] \quad (2.13)$$

We have used Lemma 2.1.1 and set:

$$\Omega^{(2)}_{N,t}[\,\cdot\,] = \sum_{\sigma,\tau\in\Sigma_N}\frac{e^{-\beta H_N(\sigma;t)}}{Z_N(t)}\frac{e^{-\beta H_N(\tau;t)}}{Z_N(t)}(\cdot) \quad (2.14)$$

Now, thanks to the convexity of the square:

$$q_N(\sigma,\tau) = \frac{N_1}{N}q_{N,1}(\sigma,\tau) + \frac{N_2}{N}q_{N,2}(\sigma,\tau) \quad \Rightarrow \quad q^2_N(\sigma,\tau) \le \frac{N_1}{N}q^2_{N,1}(\sigma,\tau) + \frac{N_2}{N}q^2_{N,2}(\sigma,\tau) \quad (2.15)$$

The result follows immediately, after reinserting the above inequality in the first derivitave of the interpolating pressure. □

2.1.2 Normalization stability

Unfortunately, we are not able yet to compute the finite size corrections to the pressure of the model, thus we cannot evaluate the effect that a change in normalization could have on the thermodynamic limit. However, one can prove the following "normalization stability" theorem.

Theorem 2.1.3 (Normalization stability of the thermodynamic limit, SK)**.** *The SK hamiltonian* (2.1) *and the following one:*

$$-\tilde{H}^{SK}_N(\sigma;J,h) = K_{N,c}(\sigma;J) + h\sum_{i=1}^N \sigma_i \quad (2.16)$$

$$K_{N,c}(\sigma;J) = \frac{1}{\sqrt{N+c}}\sum_{i,j=1}^N J_{ij}\sigma_i\sigma_j \quad (2.17)$$

where $c \in \mathbb{R}$, induce the same quenched pressure per particle in the thermodynamic limit:

$$\lim_{N\to\infty} (p_N^{SK} - \tilde{p}_N^{SK}) = 0 \tag{2.18}$$

$$p_N^{SK} = \mathbb{E}\frac{1}{N}\log\sum_{\sigma\in\Sigma_N} e^{-\beta H_N^{SK}(\sigma)} \tag{2.19}$$

$$\tilde{p}_N^{SK} = \mathbb{E}\frac{1}{N}\log\sum_{\sigma\in\Sigma_N} e^{-\beta \tilde{H}_N^{SK}(\sigma)} \tag{2.20}$$

Proof. We follow the same steps of the previous proof. First, we build up the interpolating Hamiltonian and its pressure:

$$H_\sigma(t) = \frac{-1}{\sqrt{N+tc}}\sum_{i,j=1}^N J_{ij}\sigma_i\sigma_j - h\sum_{i=1}^N \sigma_i = -K_\sigma(t) - h\sum_{i\in I_N}\sigma_i \tag{2.21}$$

$$K_\sigma(t) = \frac{1}{\sqrt{N+tc}}\sum_{i,j=1}^N J_{ij}\sigma_i\sigma_j \tag{2.22}$$

$$P(t) = \mathbb{E}\log\underbrace{\sum_{\sigma\in\Sigma_N} e^{-\beta H_\sigma(t)}}_{Z_N(t)} = \mathbb{E}\log\sum_{\sigma\in\Sigma_N} W_N(\sigma;\beta,h)\exp\left(\beta K_\sigma(t)\right) \tag{2.23}$$

$$P(0) = P_N^{SK} \qquad P(1) = \tilde{P}_N^{SK} \tag{2.24}$$

The derivative of the interpolating interaction hamiltonian is:

$$\frac{dK_\sigma(t)}{dt} = -\frac{c}{2(N+tc)^{3/2}}\sum_{i,j=1}^N J_{ij}\sigma_i\sigma_j = -\frac{c}{2(N+tc)}K_\sigma(t) \tag{2.25}$$

Turning back to the pressure:

$$P'(t) = -\frac{\beta c}{2(N+tc)}\sum_{\sigma\in\Sigma_N}\mathbb{E}\left[\frac{W_N(\sigma;\beta,h)\exp\left(\beta K_\sigma(t)\right)}{Z_N(t)}K_\sigma(t)\right] \tag{2.26}$$

Now we use the integration by parts (Lemma 2.1.1).

$$P'(t) = \frac{-\beta^2 c}{2(N+tc)}\sum_{\sigma,\tau\in\Sigma_N} C_{\sigma\tau}(t)\mathbb{E}\left[\frac{e^{-\beta H_\sigma(t)}\delta_{\sigma\tau}}{Z(t)} - \frac{e^{-\beta H_\sigma(t)-\beta H_\tau(t)}}{Z^2(t)}\right] =$$

$$= \frac{-\beta^2 c}{2(N+tc)}\mathbb{E}\Omega_{N,t}^{(2)}\left[C_{\sigma\sigma}(t) - C_{\sigma\tau}(t)\right] \tag{2.27}$$

The covariance $C_{\sigma\tau}(t)$ has a simple explicit expression with the selected interpolating Hamiltonian, thanks to the fact that the J_{ij} are i.i.d., centered and with unit variance:

$$C_{\sigma\tau}(t) = \frac{1}{N+tc} \sum_{i,j,k,l \in I_N} \sigma_i \sigma_j \tau_k \tau_l \mathbb{E}[J_{ij} J_{kl}] =$$
$$= \frac{N^2}{N+tc} \left(\frac{1}{N} \sum_{i \in I_N} \sigma_i \tau_i \right)^2 = \frac{N^2}{N+tc} q_N^2(\sigma, \tau) \quad (2.28)$$

Inserting it in the expression for $P'(t)$ we finally get:

$$P'(t) = \frac{-\beta^2 N^2 c}{2(N+tc)^2} \left[1 - \mathbb{E}\Omega^{(2)}_{N,t}[q_N^2(\sigma,\tau)] \right] \quad (2.29)$$

As seen in the proof for the Curie-Weiss deterministic model it is the integral of $P'(t)/N$ to determine whether the difference between the two pressures vanishes in the thermodynamic limit. Thus we focus on:

$$\left| \frac{1}{N} \int_0^1 dt \, P'(t) \right| \leq \frac{\beta^2 N |c|}{2} \int_0^1 \frac{dt}{(N+tc)^2} = \frac{\beta^2 N |c|}{2N(N+c)} \longrightarrow 0 \quad (2.30)$$

when $N \to \infty$. Hence:

$$\tilde{p}_N^{SK} - p_N^{SK} = \frac{1}{N} \int_0^1 dt \, P'(t) \longrightarrow 0 \quad (2.31)$$

□

2.2 Replica symmetric solution

As already specified in the introduction to this chapter, the so called *replica symmetric solution* is only an approximation of the real pressure of the SK model in the thermodynamic limit. This *ansatz* is named after the technique used to obtain it, the replica method. We will not show it here in detail, the reader can find it explained in [10].

However one can rigorously prove the following upper bound, which will be very important to discuss some properties in certain regimes in the phase space (β, h) of our system in Chapter 5.

Theorem 2.2.1 (Replica symmetric bound). *The pressure of the SK model is bounded from above by the replica symmetric pressure. Precisely:*

$$p_N^{SK}(\beta, h) \leq p_{RS}(\beta, h; q) \tag{2.32}$$

$$p_{RS}(\beta, h; q) = \frac{\beta^2}{2}(1-q)^2 + \log 2 + \mathbb{E}_z \log\cosh\left[\beta\left(h + \sqrt{2q}z\right)\right] \tag{2.33}$$

where $\mathbb{E}_z$ denotes the expectation w.r.t. the standard gaussian variable z.

The consistency equation deriving from the replica symmetric pressure, optimizing w.r.t. q, is:

$$\bar{q} = \mathbb{E}_z \tanh^2\left[\beta\left(h + \sqrt{2\bar{q}}z\right)\right] \tag{2.34}$$

Proof. The key is again the interpolation technique. Consider:

$$\tilde{H}_N(\sigma, \tilde{J}; q) = -\sqrt{2q}\sum_{i=1}^N \tilde{J}_i\sigma_i \quad \tilde{J}_i \overset{\text{iid}}{\sim} \mathcal{N}(0,1) \tag{2.35}$$

$$\tilde{p}_N = \frac{1}{N}\mathbb{E}\log\sum_\sigma W_N(\beta, h; \sigma)e^{-\beta\tilde{H}_N(\sigma)} = \mathbb{E}\log 2\cosh\left[\beta\left(h + z\sqrt{2q}\right)\right] \tag{2.36}$$

Define then the interpolation hamiltonian:

$$-H_\sigma(t) = \sqrt{t}K_N(\sigma; J) + \sqrt{1-t}\sqrt{2q}\sum_{i=1}^N \tilde{J}_i\sigma_i + h\sum_{i=1}^N \sigma_i \tag{2.37}$$

Let us now compare the covariances of the two gaussian families:

$$\mathbb{E}[K_N(\sigma)K_N(\tau)] = Nq_N^2(\sigma, \tau) \tag{2.38}$$

$$\mathbb{E}[\tilde{H}_N(\sigma, \tilde{J}; q)\tilde{H}_N(\tau, \tilde{J}; q)] = N2qq_N(\sigma, \tau) \tag{2.39}$$

Keeping that in mind, we are ready to compute the first derviative of the interpolating pressure:

$$\begin{aligned} p'_N(t) = \frac{\beta}{N}\mathbb{E}\omega_{N,t}\left[\frac{1}{2\sqrt{t}}K_\sigma - \frac{1}{2\sqrt{1-t}}\tilde{H}_\sigma\right] &= \\ = \frac{\beta^2}{2}\mathbb{E}\omega_{N,t}\left[1 - q_N^2(\sigma,\tau) - 2q + 2qq_N(\sigma,\tau)\right] &= \\ = -\frac{\beta^2}{2}\mathbb{E}\omega_{N,t}[(q_{\sigma\tau} - q)^2] + \frac{\beta^2}{2}(1-q)^2 &\leq \frac{\beta^2}{2}(1-q)^2 \end{aligned} \tag{2.40}$$

where we have adopted a more convenient notation as done in the previous sections and used the gaussian integration by parts. Hence, by the theorem of integral calculus:

$$p_N^{SK} = \tilde{p}_N + \int_0^1 dt\, p_N'(t) \leq \tilde{p}_N + \frac{\beta^2}{2}(1-q)^2 \tag{2.41}$$

After the insertion of $\tilde{p}_N$ it is clear that the r.h.s. is exactly the p_{RS} functional.

The consistency equation is simply obtained deriving w.r.t. q the replica symmetric pressure and using integration by parts. $\square$

2.3 ROSt-cavity perspective

2.3.1 Poisson Point Processes: basic properties

We recall here only a few basic notions concerning PPP's useful for our purposes, for more details see [10] and [18].

Definition 2.3.1. Let $E \subseteq \mathbb{R}$, $(X_i)_{i\geq 1}$ be a sequence of points randomly thrown on it and A an interval in E. The point process is said to be a PPP of intensity measure μ if and only if:

- $N(A) = \sum_i \mathbb{1}_A(X_i)$ is $\sim Poisson(\mu(A))$
- $A_1, \ldots, A_n$ disjoint $\Rightarrow N(A_1), \ldots, N(A_n)$ are independent r.v.

In order to build Ruelle cascades to represent the Parisi functional, on needs to know how to generate Poisson processes with a given intensity measure μ.

Proposition 2.3.1. *Consider $E \subseteq \mathbb{R}$ and an intensity measure μ with $\mu(E) < \infty$. The process obtained sampling the number of points from $Poisson(\mu(E))$ and then distributing them identically and independently according to:*

$$\tilde{\mu}(dx) = \frac{\mu(dx)}{\mu(E)} \tag{2.42}$$

is a PPP.

Proof. Let us compute the characteristic function of the number of points in a set $A \subseteq E$. Thanks to the total probability formula we get:

$$\mathbb{E}\left[e^{-tN(A)}\right] = \sum_{n=0}^{\infty} \mathbb{E}\left[e^{-t\sum_i \mathbb{1}_A(X_i)} \mid N(E) = n\right] \mathbb{P}(N(E) = n) =$$

$$\sum_{n=0}^{\infty} \mathbb{E}^n\left[e^{-t\mathbb{1}_A(X_i)} \mid N(E) = n\right] \mathbb{P}(N(E) = n) =$$

$$= \sum_{n=0}^{\infty} \left[\frac{e^{-t}\mu(A) + \mu(E) - \mu(A)}{\mu(E)}\right]^n \frac{\mu^n(E)}{n!} e^{-\mu(E)} = e^{\mu(A)(e^{-t}-1)} \tag{2.43}$$

which is the characteristic function of a Poisson process. The indipendence of $N(A_1), \ldots, N(A_n)$ if $A_1, \ldots, A_n$ are disjoint can be proved always through the generating functionals. □

PPP's enjoy a transformation property. We will only report the statement here, the proof is elementary and can be found in [10].

Proposition 2.3.2 (Transformation of a PPP)**.** *Let $N, (X_i)_{i\geq 1}$ be a PPP on E with intensity μ. Moreover, consider a map $f : E \longrightarrow E'$ such that A' bounded in E' $\Rightarrow$ $f^{-1}(A')$ bounded in E. Then $N \circ f^{-1}, f(X_i)$ is a PPP on E' with intensity $\mu \circ f^{-1}$.*

For our purposes, we are interested in exponential intensity PPP *i.e.* with $\mu(dx) = e^{-x}dx$. However $\mu(\mathbb{R}) = \infty$, thus one has to put a cut off at a certain $c < 0$ obtaining a modified measure:

$$\mu_c(dx) = e^{-x}\theta(x - c)dx \quad \Rightarrow \quad \mu_c(\mathbb{R}) = e^{-c} \tag{2.44}$$

$$\tilde{\mu}_c(dx) = \theta(x - c)e^{-(x-c)}dx \tag{2.45}$$

then, letting $c \to -\infty$ we recover the correct PPP over the entire real line.

Let us now write down the invariance property of an exponential intensity PPP that, together with the transformation properties, will allow us to represent the Parisi functional. See [10] again for a brief proof.

Theorem 2.3.3 (Invariance property of exponential intensity PPP)**.** *Let $(X_i)_{i\geq 1}$ be the points of a PPP with intensity $\mu(dx) = e^{-x}dx$ and $(U_i)_{i\geq 1}$ a sequence of i.i.d. variables such that $\mathbb{E}[e^{mU_i}] < \infty$ for a fixed $m \in (0, 1)$. Then:*

$$e^{X_i/m}e^{U_i} \overset{\mathrm{D}}{=} \mathbb{E}[e^{mU_i}]^{1/m}e^{X_i/m} \tag{2.46}$$

2.3.2 Ruelle Probability Cascades (RPC)

We can generate a Ruelle cascade starting from a PPP with exponential intensity, so that we can use the invariance property that will be crucial to reproduce the nested nature of the Parisi *ansatz*.

To begin with, consider a PPP with intensity $e^{-x}dx$ whose points are $\{X_\alpha\}_{\alpha\in\mathbb{N}}$ and a non deacreasing sequence of real numbers $0 = m_0 \leq m_1 \leq m_2 \leq \cdots \leq m_k = 1$. Through the transformation property, we can define another PPP with intensity $m_1 y^{-m_1-1}dy$, namely $Y_{\alpha_1} = \exp(X_{\alpha_1}/m_1)$. Then for every index $\alpha_1 \in \mathbb{N}$ we generate another independent PPP with intensity $m_2 y^{-m_2-1}dy$. This can be done again trough an exponential PPP $X_{\alpha_1\alpha_2}$, $Y_{\alpha_1\alpha_2} = \exp(X_{\alpha_1\alpha_2}/m_2)$. We can go on iteratively up to the k-th step.

When this generation stops, we end up with a structure of indices similar to a cascade or a tree, with a root conventionally set in $\mathbb{N}^0 = \{\emptyset\}$. To each internal vertex of these trees (excluding the leaves labeled by k natural numbers) there is an independent PPP associated to it.

Define now the following random weights:

$$\bar{\xi}_\alpha = Y_{\alpha_1} Y_{\alpha_1\alpha_2} \dots Y_{\alpha_1\dots\alpha_k} \quad \alpha = (\alpha_1, \dots, \alpha_k) \tag{2.47}$$

Thanks to the already mentioned invariance property of PPPs the following lemma holds. For a complete proof see [18].

Lemma 2.3.4. *With probability one, the sum* $\sum_\alpha \bar{\xi}_\alpha < \infty$ *and, hence, the sequence* (2.47) *is well defined.*

As we shall see later, this allows us to write a ROSt with random weights build on RPC.

2.3.3 Random Overlap Structures (ROSt) and cavity functional

Here we will provide the extended variational principle formulated by Aizenman, Sims and Starr in [2] and [3]. This "scheme" is proved to contain the Guerra interpolation scheme, though non trivial arguments are needed to show it. For the moment we will consider a hamiltonian normalized with a $1/\sqrt{2N}$ factor for convenience. The results are still valid for (2.1).

Imagine to have a system of N spins $\alpha \in \Sigma_N$ and to add a single spin, denoted by $\hat{\sigma}$. We can rewrite the quenched pressure as a Cesaro summation:

$$p_N^{SK} = \frac{1}{N}\mathbb{E}\log Z_N = \frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}\log\left(\frac{Z_{n+1}}{Z_n}\right) \tag{2.48}$$

$$\lim_{N\to\infty} p_N^{SK} = \lim_{N\to\infty}\mathbb{E}\log\left(\frac{Z_{N+1}}{Z_N}\right) \tag{2.49}$$

Hence we can focus only on the difference of the two pressures, that is nothing but the increment after the addition of $\hat{\sigma}$ to the system.

$$\mathbb{E}\log\left[\frac{\sum_{\alpha,\hat{\sigma}}\exp\left(-H_{N+1}^{SK}(\alpha,\hat{\sigma})\right)}{\sum_\alpha \exp\left(-H_N^{SK}(\alpha)\right)}\right] \tag{2.50}$$

Thanks to the fact that we have an expectation over the gaussian disorder we can replace H_N^{SK} with the following equality in distribution:

$$\begin{aligned} -H_N^{SK}(\alpha; J) = \frac{1}{\sqrt{2N}}\sum_{i,j=1}^N J_{ij}\alpha_i\alpha_j \overset{\mathrm{D}}{=} \sqrt{\frac{1}{2}}\underbrace{\frac{1}{\sqrt{N(N+1)}}\sum_{i,j=1}^N \tilde{J}_{ij}\alpha_i\alpha_j +}_{K(\alpha)} \\ + \frac{1}{\sqrt{2(N+1)}}\sum_{i,j=1}^N J_{ij}\alpha_i\alpha_j \end{aligned} \tag{2.51}$$

where $\tilde{J}_{ij}$ are independent gaussian centered r.v.. We split then H_{N+1}^{SK} in three parts:

$$\begin{aligned} -H_{N+1}^{SK}(\alpha,\hat{\sigma}) = \frac{1}{\sqrt{2(N+1)}}\sum_{i,j=1}^N J_{ij}\alpha_i\alpha_j + \frac{1}{\sqrt{(N+1)}}\sum_{i=1}^N J_{i,N+1}\alpha_i\hat{\sigma} + \\ + \frac{1}{\sqrt{2(N+1)}}J_{N+1,N+1} \end{aligned} \tag{2.52}$$

The last term is irrelevant in the thermodynamic limit. We have neglected the magnetic field term up to now, since it is deterministic. Let us reintroduce it. We can recast the incremental pressure up to irrelevant terms as follows:

$$\mathbb{E}\log\left(\frac{Z_{N+1}}{Z_N}\right)=\mathbb{E}\log\left[\frac{\sum_{\alpha,\hat{\sigma}}\xi_\alpha\exp\left(\beta(\eta_\alpha+h)\hat{\sigma}\right)}{\sum_\alpha\xi_\alpha\exp\left(\frac{\beta}{\sqrt{2}}K(\alpha)\right)}\right] \tag{2.53}$$

$$K(\alpha)=\frac{1}{\sqrt{N(N+1)}}\sum_{i,j=1}^{N}\tilde{J}_{ij}\alpha_i\alpha_j \tag{2.54}$$

$$\xi_\alpha=\exp\left(\frac{\beta}{\sqrt{2(N+1)}}\sum_{i,j=1}^{N}J_{ij}\alpha_i\alpha_j+\beta h\sum_{i=1}^{N}\alpha_i\right) \tag{2.55}$$

$$\eta_\alpha=\frac{1}{\sqrt{N+1}}\sum_{i=1}^{N}J_{i,N+1}\alpha_i \tag{2.56}$$

One could generalize to M spins σ introduced in the system with the same arguments, obtaining an object similar to the cavity functional (see below). Inspired by the previous construction, we give the following definitions. The reader will notice that some notations will not change, in order to make some associations and identifications easier.

Definition 2.3.2 (ROSt)**.** Let $\{\xi_\alpha\}$ be random wights with law μ and $\{p_{\alpha,\alpha'}\}$ a covariance matrix. The couple $r=(p,\xi)$ is a ROSt if:

- $\xi_\alpha\geq 0$ and $\sum_\alpha\xi_\alpha<\infty$ μ-a.s.;
- the matrix p is positive definite;
- $p_{\alpha,\alpha}=1$.

Consider now two gaussian fields, fully characterized by their covariances:

$$\mathbb{E}[\eta_{j,\alpha}\eta_{j'\alpha'}]=\delta_{jj'}p_{\alpha\alpha'} \tag{2.57}$$

$$\mathbb{E}[K_\alpha K'_\alpha]=p^2_{\alpha\alpha'} \tag{2.58}$$

Remark 2.3.1. Notiche that the fields defined in (2.54) and (2.56) have the following covariances:

$$\mathbb{E}[\eta_\alpha\eta_{\alpha'}]=\frac{N}{N+1}\bar{p}_{\alpha\alpha'}\quad\bar{p}_{\alpha\alpha'}=\frac{1}{N}\sum_{i=1}^{N}\alpha_i\alpha'_i \tag{2.59}$$

$$\mathbb{E}[K(\alpha)K(\alpha')]=\frac{N}{N+1}\bar{p}^2_{\alpha\alpha'} \tag{2.60}$$

It is easy to see that in the thermodynamic limit, with an appropriate identification, these two fields coincide with those defined to introduce the cavity functional (2.57), (2.58).

Definition 2.3.3 (Cavity functional)**.** Let $r = (\xi, p)$ be a ROSt and consider a system of M spins, $\sigma \in \Sigma_M$, added to the original system with N spins, $\alpha \in \Sigma_N$. The cavity functional is:

$$G_{r,M}(\beta, h) = \frac{1}{M}\mathbb{E}\log\left[\frac{\sum_{\alpha,\sigma}\xi_\alpha \exp\left(\beta\left(\sum_{j=1}^{M}\eta_{j,\alpha} + h\right)\sigma_j\right)}{\sum_\alpha \xi_\alpha \exp\left(\beta\sqrt{\frac{M}{2}}K(\alpha)\right)}\right] \tag{2.61}$$

where $\eta_{j,\alpha}$ and K_α have been defined in (2.57) and (2.58) respectively.

Now we proceed with the proof of the following important proposition that will provide the upper bound for the pressure of the SK model.

Proposition 2.3.5. $\forall\, M \in \mathbb{N}$ *and for any ROSt* r *we have:*

$$p_M^{SK}(\beta, h) \leq G_{r,M}(\beta, h) \tag{2.62}$$

Proof. Let us write the l.h.s. more explicitly.

$$\begin{aligned} p_M^{SK}(\beta, h) &= \frac{1}{M}\mathbb{E}\log\left[\sum_\sigma \exp\left(\beta\left(\frac{1}{\sqrt{2M}}\sum_{i,j=1}^{M} J_{ij}\sigma_i\sigma_j + h\sum_{j=1}^{M}\sigma_j\right)\right)\right] = \\ &= \frac{1}{M}\mathbb{E}\log\left[\sum_\sigma \exp\left(\frac{\beta}{\sqrt{2M}}\sum_{i,j=1}^{M} J_{ij}\sigma_i\sigma_j\right) W_M(\beta, h; \sigma)\right] \end{aligned} \tag{2.63}$$

where W_M is a deterministic wight. The statement can thus be recast in the following form:

$$\begin{aligned} &\mathbb{E}\log\left[\sum_{\alpha\sigma}\xi_\alpha \exp\left(\beta\overbrace{\left(\frac{1}{\sqrt{2M}}\sum_{i,j=1}^{M} J_{ij}\sigma_i\sigma_j - \sqrt{\frac{M}{2}}K(\alpha)\right)}^{A_{\sigma,\alpha}}\right) W_M(\beta, h; \sigma)\right] \leq \\ &\qquad \leq \mathbb{E}\log\left[\sum_{\alpha,\sigma}\xi_\alpha \exp\left(\beta\underbrace{\sum_{j=1}^{M}\eta_{j,\alpha}\sigma_j}_{B_{\sigma\alpha}}\right) W_M(\beta, h; \sigma)\right] \end{aligned} \tag{2.64}$$

The two disorders A and B are manifestly independent and characterized by the following covariances:

$$\mathbb{E}[A_{\sigma\alpha}A_{\sigma'\alpha'}] = \frac{M}{2}q_M^2(\sigma,\sigma') + \frac{M}{2}p_{\alpha\alpha'}^2 \tag{2.65}$$
$$\mathbb{E}[B_{\sigma\alpha}B_{\sigma'\alpha'}] = Mq_M(\sigma,\sigma')p_{\alpha\alpha'} \tag{2.66}$$
$$\mathbb{E}[A_{\sigma\alpha}A_{\sigma\alpha}] = \mathbb{E}[B_{\sigma\alpha}B_{\sigma\alpha}] \tag{2.67}$$
$$\mathbb{E}[A_{\sigma\alpha}A_{\sigma'\alpha'}] \geq \mathbb{E}[B_{\sigma\alpha}B_{\sigma'\alpha'}] \tag{2.68}$$

The last inequality is due to the convexity of the square. The result now follows from the Comparison of Gaussian Families Theorem in [10] (Theorem 3.46). $\square$

Let us now see if we manage to find also a lower bound in terms of the cavity functional

Lemma 2.3.6. *Let Q_N be a sequence of real numbers.The following limiting inequality holds:*

$$\liminf_{N\to\infty}\frac{Q_N}{N} \geq \liminf_{N\to\infty}\frac{Q_{N+M}-Q_N}{M} \tag{2.69}$$

Proof. Consider three positive integers n, M, N.

$$\frac{Q_{nM+N}-Q_N}{nM+N} = \frac{\sum_{j=0}^{n-1}(Q_{(j+1)M+N}-Q_{jM+N})}{nM+N} \geq \frac{n\inf_{k\geq N}(Q_{k+M}-Q_k)}{nM+N} \tag{2.70}$$

It suffices to take the $\liminf_{n\to\infty}$ of both sides and then the $\sup_N$ of the r.h.s. to optimize, and te result is proved. $\square$

Theorem 2.3.7 (Aizenman-Sims-Starr extended variational principle)**.** *The pressure of the SK model in the thermodynamic limit is:*

$$p^{SK}(\beta,h) = \lim_{M\to\infty}\inf_r G_{r,M}(\beta,h) \tag{2.71}$$

Proof. Optimizing with respect to the possible ROSts r in Proposition 2.3.5 and then sending $M\to\infty$ we immediately have the upper bound.

The lower bound is obtained by means of Lemma 2.3.6. We just need to exhibit a ROSt r such that:

$$\liminf_{N\to\infty}\frac{P_{N+M}-P_N}{M} = G_{r,M}(\beta,h) \tag{2.72}$$

This can be done by slightly modifying the definitions (2.54) and (2.56). In fact, as already said, if we imagine to add M spins σ_i, instead of one, we would get the following fields and weights:

$$K_M(\alpha) = \frac{1}{\sqrt{N(N+M)}} \sum_{i,j=1}^{N} J_{ij}\alpha_i\alpha_j \tag{2.73}$$

$$\eta_{j,\alpha} = \frac{1}{\sqrt{N+M}} \sum_{i=1}^{N} J_{ij}\alpha_i \tag{2.74}$$

$$\xi_\alpha = \exp\left(\frac{\beta}{\sqrt{2(N+M)}} \sum_{i,j=1}^{N} J_{ij}\alpha_i\alpha_j + \beta h \sum_{i=1}^{N} \alpha_i\right) \tag{2.75}$$

with covariances:

$$\mathbb{E}[\eta_{j,\alpha}\eta_{j',\alpha'}] = \frac{N}{N+M}\delta_{jj'}\bar{p}_{\alpha\alpha'} \quad \bar{p}_{\alpha\alpha'} = \frac{1}{N}\sum_{i=1}^{N}\alpha_i\alpha'_i \tag{2.76}$$

$$\mathbb{E}[K_M(\alpha)K_M(\alpha')] = \frac{N}{N+M}\bar{p}^2_{\alpha\alpha'} \tag{2.77}$$

It is easy to see that in the $\liminf_{N\to\infty}$ the previous covariances reproduce those of a cavity functional fields. An extra field would appear in the argument of the exponential, *i.e.*:

$$U(\sigma) = \frac{1}{\sqrt{2(N+M)}} \sum_{i,j=1}^{M} J_{ij}\sigma_i\sigma_j \tag{2.78}$$

that is irrelevant if we perform before the limit in N.

Finally, to sum up, we have proved that:

$$\lim_{N\to\infty} p_N^{SK} \geq \liminf_{N\to\infty} p_N^{SK} \geq \inf_r G_{r,M} \tag{2.79}$$

Then, sending $M \to \infty$ we get the result. □

2.4 Replica symmetry breaking solution

In order to introduce the Parisi formula, one needs a certain number of definitions. To begin with let us introduce two non decreasing sequences:

$$0 = m_0 \leq m_1 \leq \cdots \leq m_k \leq m_{k+1} = 1 \tag{2.80}$$

$$0 = q_0 \leq q_1 \leq \cdots \leq q_{k-1} \leq q_k = 1 \tag{2.81}$$

where k is a positive integer.

We can associate a cumulative probability $x(q)$ of a discrete variable to the triple $(k, \mathbf{m}, \mathbf{q})$, that will play the role of our order parameter as follows:

$$x(q) = \sum_{l=0}^{k} (m_{l+1} - m_l)\theta(q - q_l) \tag{2.82}$$

where the Heaviside step function is taken to be continuous from the right.

Let us now introduce the recursive definition:

$$Z_{l-1}^{m_l} = \mathbb{E}_l \left(Z_l^{m_l} \right) \quad \mathbb{E}_l[\cdot] = \int_{\mathbb{R}} \frac{d\eta_l}{\sqrt{2\pi}} e^{-\eta_l^2/2} (\cdot) \tag{2.83}$$

$$Z_k = \cosh\left[\beta\left(h + \sum_{l=1}^{k} \eta_l \sqrt{q - q_{l-1}}\right)\right] \quad \eta_l \overset{\text{iid}}{\sim} \mathcal{N}(0,1) \tag{2.84}$$

The Replica Symmetry Breaking (RSB) *ansatz* is defined as follows.

Definition 2.4.1 (Parisi Functional)**.** Given the triple $k, \mathbf{m}, \mathbf{q}$ as before, the Parisi functional is:

$$\mathcal{P}(x; \beta, h) = \log 2 + \log Z_0 - \frac{\beta^2}{2} \int_0^1 q x(q)\, dq \tag{2.85}$$

We are finally ready to write the main statement, rigorously proved by Guerra (2003) and Talagrand (2006).

Theorem 2.4.1 (Pressure particle of the SK model)**.** *The pressure of the SK model in the thermodynamic limit is given by the infimum over all the possible choices of the triples $k, \mathbf{q}, \mathbf{m}$ of the Parisi functional, i.e.:*

$$p^{SK}(\beta, h) = \inf_{x(q)} \mathcal{P}(x(q); \beta, h) \tag{2.86}$$

Remark 2.4.1. Notice that the infimum could actually be taken over the possible non decreasing functions $x(q) : [0,1] \longrightarrow [0,1]$, as proved by Guerra in [14] by a continuity argument. Thanks to this fact, and for convenience, it will be sufficient to deal with discrete distributions ([26]).

2.4.1 Representation of the Parisi functional via RPC

Let us begin with the following definition, that will turn out to be very useful in the proof of the theorem below.

Definition 2.4.2 (Ultrametric structure of an overlap kernel)**.** The overlap kernel $\bar{p}_{\alpha\alpha'}$ is said to be ultrametric if:

$$\bar{p}_{\alpha\alpha'} \begin{cases} q_0 & \text{if } \alpha_1 \neq \alpha'_1 \\ q_1 & \text{if } \alpha_1 = \alpha'_1 \text{ and } \alpha_2 \neq \alpha'_2 \\ \vdots \\ q_k = 1 & \text{if } \alpha_i = \alpha'_i \ \forall\, i = 1, \dots, k \end{cases} \tag{2.87}$$

where $(q_l)_{l \leq k}$ is a non decreasing sequence.

Theorem 2.4.2. *Consider a ROSt $\bar{r} = (\bar{\xi}, \bar{p})$ where $\bar{\xi}_\alpha$ are realized in terms of random weights as in* (2.47) *and $\bar{p}_{\alpha\alpha'}$ has an ultrametric structure with $\mathbf{q}$ as parameters. Then:*

$$G_{\bar{r},M}(\beta, h) = \mathcal{P}(x(q); \beta, h) \tag{2.88}$$

Proof. First we need to define some fields that will enter the cavity functional with the correct covariances. It is really easy to check that the following gaussian r.v.:

$$\bar{\eta}_{j,\alpha} = \sqrt{q_1 - q_0} J_{j,\alpha_1} + \sqrt{q_2 - q_1} J_{j,\alpha_1\alpha_2} + \cdots + \sqrt{q_k - q_{k-1}} J_{j,\alpha_1\dots\alpha_k} \quad j = 1, \dots, M \tag{2.89}$$

$$\bar{K}_\alpha = \sqrt{q_1^2 - q_0^2} \tilde{J}_{\alpha_1} + \cdots + \sqrt{q_k^2 - q_{k-1}^2} \tilde{J}_{\alpha_1\dots\alpha_k} \tag{2.90}$$

where the J's and $\tilde{J}$'s are $\overset{\text{iid}}{\sim} \mathcal{N}(0,1)$, are still gaussian r.v.. In addition to that, they satisfy (2.57) and (2.58).

Now, recalling the invariance property (2.46), and applying it for k steps, we can write:

$$\begin{aligned} &\frac{1}{M} \mathbb{E} \log \left[\sum_{\sigma,\alpha} \bar{\xi}_\alpha \exp \left(\beta \sum_{j=1}^{M} \left(\bar{\eta}_{j,\alpha} + h \right) \sigma_j \right) \right] = \\ &\qquad = \frac{1}{M} \mathbb{E} \log \left[\sum_{\sigma,\alpha} \bar{\xi}_\alpha \prod_{j=1}^{M} e^{\log 2 \cosh \beta \left(\sqrt{q_1 - q_0} J_{j,\alpha_1} + \cdots + \sqrt{q_k - q_{k-1}} J_{j,\alpha_1\dots\alpha_k} + h \right)} \right] = \\ &\qquad = \frac{1}{M} \mathbb{E} \log \left[\sum_{\sigma,\alpha} \bar{\xi}_\alpha \prod_{j=1}^{M} \mathbb{E}_k^{1/m_k} \left(e^{m_k \log 2 \cosh \beta \left(\sqrt{q_1 - q_0} J_{j,\alpha_1} + \cdots + \sqrt{q_k - q_{k-1}} \eta_k + h \right)} \right) \right] = \\ &\qquad = \cdots = \log 2 + \log Z_0 + \mathbb{E} \log \sum_{\alpha} \bar{\xi}_\alpha \end{aligned} \tag{2.91}$$

The same trick can be used for the denominator of the cavity functional,:

$$\begin{aligned}
&\frac{1}{M}\mathbb{E}\log\left[\sum_\alpha \bar{\xi}_\alpha \exp\left(\beta\sqrt{\frac{M}{2}}\bar{K}_\alpha\right)\right] = \\
&\qquad = \frac{1}{M}\mathbb{E}\log\left[\sum_\alpha \bar{\xi}_\alpha e^{\beta\sqrt{\frac{M}{2}}\left(\sqrt{q_1^2-q_0^2}\tilde{J}_{\alpha_1}+\dots+\sqrt{q_k^2-q_{k-1}^2}\tilde{J}_{\alpha_1\dots\alpha_k}\right)}\right] = \\
&\qquad = \frac{1}{M}\mathbb{E}\log\left[\sum_\alpha \bar{\xi}_\alpha \mathbb{E}_k^{1/m_k} e^{m_k\beta\sqrt{\frac{M}{2}}\left(\sqrt{q_1^2-q_0^2}\tilde{J}_{\alpha_1}+\dots+\sqrt{q_k^2-q_{k-1}^2}\eta_k\right)}\right] \qquad (2.92)
\end{aligned}$$

Thanks to Hubbard-Stratonovič transform, *i.e.* the computation of the generating function of a gaussian process, one immediately gets:

$$\begin{aligned}
\frac{1}{M}\mathbb{E}\log\left[\sum_\alpha \bar{\xi}_\alpha \exp\left(\beta\sqrt{\frac{M}{2}}\bar{K}_\alpha\right)\right] &= \mathbb{E}\log\sum_\alpha \bar{\xi}_\alpha + \frac{\beta^2}{4}\sum_{l=1}^k m_l(q_l^2-q_{l-1}^2) = \\
&= \mathbb{E}\log\sum_\alpha \bar{\xi}_\alpha + \frac{\beta^2}{2}\int_0^1 dq\, x(q)q \qquad (2.93)
\end{aligned}$$

with a piecwise constant $x(q)$ as described in the definition of the Parisi functional.

□

Corollary 2.4.3 (Upper bound for the SK pressure)**.**

$$p_N^{SK}(\beta,h) \le \mathcal{P}(x(q);\beta,h) \qquad (2.94)$$

Proof. It follows immediately from the previous theorem and from Proposition 2.3.5.

□

2.4.2 Lower bound of the SK model pressure

We have clearly seen that Ruelle cascades play a central role in the SK model. In particular, up to now, we have only tackled the problem of the upper bound, with positive results, through a representation of the Parisi functional and the Aizenmann-Sims-Starr scheme. However, the lower bound is much more tricky and requires a deeper understanding of the links between ultrametricity, the *Ghirlanda-Guerra identities* (briefly discussed below) and the invariance properties of RPCs. This is why in this subsection we omit the details of the proof (contained in [18] or [26]), reporting only the main steps that will bring us to the lower bound.

RPCs and Ghirlanda Guerra identities For the sake of brevity we will introduce the notations:

$$\nu_\alpha = \frac{\bar{\xi}_\alpha}{\sum_{\alpha'} \bar{\xi}_{\alpha'}}; \quad \langle \cdot \rangle = \frac{\sum_\alpha \bar{\xi}_\alpha(\cdot)}{\sum_{\alpha'} \bar{\xi}_{\alpha'}} = \sum_\alpha \nu_\alpha(\cdot) \tag{2.95}$$

with $\bar{\xi}_\alpha$ as previously defined. Notice that $\langle \cdot \rangle$ is a random measure, with respect to which functions of the overlaps may be averaged. For the precise meaning of it see [18] (eqs. (2.61), (2.70)). In his monograph, D. Panchenko proves the following result (see Lemma 2.5 and Theorem 2.10 again in [18]):

Theorem 2.4.4 (Ghirlanda-Guerra identities, RPC)**.** *Let f be a function of the overlaps $q_{ll'} = \sigma^l \cdot \sigma^{l'}$ where $\sigma^l, \sigma^{l'}$ are i.i.d. sampled according to $\langle \cdot \rangle$, and $n \geq 1$ an integer. Then, for any function $\psi : \mathbb{R} \longrightarrow \mathbb{R}$ the following Ghirlanda-Guerra identities hold:*

$$\mathbb{E}\langle f \psi(q_{1,n+1}) \rangle - \frac{1}{n}\mathbb{E}\langle f \rangle \mathbb{E}\langle \psi(q_{12}) \rangle - \frac{1}{n}\sum_{l=2}^{n} \mathbb{E}\langle f \psi(q_{1l}) \rangle = 0 \tag{2.96}$$

Furthermore, from preliminary results leading to the previous theorem, it emerges that:

$$\mathbb{E}\langle \psi(q_{12}) \rangle = \sum_{l=0}^{k} \psi(q_l)(m_{l+1} - m_l) \tag{2.97}$$

This remarkable result allows us to identify the distribution of the overlaps q_{12} in this peculiar measure $\mathbb{E}\langle \cdot \rangle$, which is nothing but (2.82).

Pay attention to the fact that we are not yet dealing with the measure induced by the SK model hamiltonian, but the previous Ghirlanda-Guerra identities will provide the missing link.

Mixed p-spins models and Ghirlanda Guerra identities In order to have results that hold for polinomials in the overlaps, we will work mixed p-spins models, defined by the hamiltonian:

$$H_N(\sigma) = \sum_{p \geq 1} \beta_p \underbrace{\frac{1}{N^{(p-1)/2}} \sum_{i_1, \ldots, i_p = 1}^{N} J_{i_1 \ldots i_p} \sigma_{i_1} \ldots \sigma_{i_p}}_{H_{N,p}} \quad J_{i_1 \ldots i_p} \overset{\text{iid}}{\sim} \mathcal{N}(0, 1) \tag{2.98}$$

Obvisouly, the covariance will contain polinomials of the overlaps. The idea used by Panchenko to force the Gibbs measure to fulfill the very same Ghirlanda-Guerra identities, is to add a perturbing hamiltonian that does not affect the thermodynamic limit. However, the derivative of the pressure will be sensible to it, thus there's hope that the identities can be properly adjusted for our purposes. The perturbing hamiltonian is:

$$g(\sigma) = \sum_{p\geq 1} \frac{x^p}{2^p} \underbrace{\frac{1}{N^{p/2}} \sum_{i_1,\dots,i_p=1}^{N} J_{i_1\dots i_p}\sigma_{i_1}\dots\sigma_{i_p}}_{q_p} \tag{2.99}$$

which is surely of a smaller order of N *w.r.t.* H_N. $x_p \in [1,2]$ are some parameters, that will be treated as uniformly distributed r.v.. This term is added to the mixed p-spins Hamiltonian together with a sequence: $H_N(\sigma)+s_N g(\sigma)$. The latter approaches 0 when $N \to \infty$ in an appropriate way, we will not write it explicitly here. The following theorem holds:

Theorem 2.4.5 (Ghirlanda-Guerra identities, mixed p-spins)**.** *If s_N satisfies condition (3.26) in [18] then:*

$$\lim_{N\to\infty} \mathbb{E}_x \left[\mathbb{E}\langle f q^p_{1,n+1}\rangle - \frac{1}{n}\mathbb{E}\langle f\rangle \mathbb{E}\langle q^p_{12}\rangle - \frac{1}{n}\sum_{l=2}^{n} \mathbb{E}\langle f q^p_{1l}\rangle \right] = 0 \tag{2.100}$$

for any $n \geq 2$ (number of replicas), $p \geq 1$ and for any function f of the overlaps.

This reminds us of the Ghirlanda-Guerra identities in [10]. The statement is basically the same as the one in Theorem 2.4.4, with $\psi(q) = q^p$. As a consequence of it we can choose a deterministic sequence $x_{p,N}$ such that the identity still holds, getting rid of the average.

Ghirlanda-Guerra identities and generic measures With the following theorem (Theorem 2.13 and Theorem 2.14 in [18]) a strong bond between the two previous paragraphs is established:

Theorem 2.4.6. *Assume that* (2.96) *holds with a generic measure* $\mathbb{E}\langle\cdot\rangle$. *Then, the distribution of the overlap matrix is uniquely determined by the distribution of* q_{12}. *Moreover, the overlap matrix is ultrametric.*

Therefore, as discussed later, when the distribution of q_{12} is discrete, it can be generated by means of Ruelle cascades, with the appropriate choice of parameters $\mathbf{q}$ and $\mathbf{m}$.

The lower bound If we follow the same steps used in the definition of the cavity functional, for the perturbed hamiltonian $H_N(\sigma) + s_N g(\sigma)$ we would get an Aizenmann-Sims-Starr representation of the type:

$$A_N(x) = \mathbb{E}\log\frac{Z'_{N+1}}{Z'_N} = \mathbb{E}\log 2\Big\langle \cosh\eta(\sigma)\Big\rangle' - \mathbb{E}\log\Big\langle \cosh K(\sigma)\Big\rangle' + o(1) \quad (2.101)$$

where $\eta(\sigma)$ and $K(\sigma)$ are two gaussian processes with the appropriate covariances. The brackets denote the expectation with respect to the Gibbs measure induced by $H'_N + s_N g$ with:

$$H'_{N_p}(\sigma) = \frac{1}{(N+1)^{(p-1)/2}}\sum_{i_1,\dots,i_p=1}^{N} J_{i_1\dots i_p}\sigma_{i_1}\dots\sigma_{i_p} \quad (2.102)$$

$$H'_N(\sigma) = \sum_{p\geq 1}\beta_p H'_{N_p}(\sigma) \quad (2.103)$$

Thanks to Lemma 2.3.6 ($M = 1$) we can state that:

$$\liminf_{N\to\infty} p_N \geq \liminf_{N\to\infty} A_N(x) \quad\Rightarrow\quad \liminf_{N\to\infty} p_N \geq \liminf_{N\to\infty}\mathbb{E}_x A_N(x) \quad (2.104)$$

It turns out that even for the measure $\mathbb{E}\langle\cdot\rangle'$ the identity (2.100) holds. Now we can replace x by x_N, eliminating the average.

It can be proved that, if $q_N(\sigma^l,\sigma^{l'})$ are the overlaps for finite N and they converge in distribution under $\mathbb{E}\langle\cdot\rangle'$ to an array $q_{ll'}$, there exists a measure $\langle\cdot\rangle$ such that the identities (2.100) containing the limiting matrix $q_{ll'}$ are satisfied $\forall p$ and $n \geq 2$. This implies also the validity of (2.96). Hence, thanks to Theorem 2.4.6, asymptotically speaking, they can be approximated in distribution with the appropriate RPC.

Let $x_n(q)$ be an order parameter converging to $x(q)$, the distribution of the limiting overlaps q. For each n there will be a corresponding Ruelle cascade determined by $x_n(q)$ itself. Let us denote it by $\langle\cdot\rangle_n$ Then, (Theorem 2.17 in [18]) we can state that the overlap matrices q_n computed with samples taken *w.r.t.* $\langle\cdot\rangle_n$ converge in distribution to the limiting matrix q. We recall that for each $x_n(q)$ we have a set of parameters that identify both a Ruelle cascade and a Parisi functional. For the moment, we denote the representation of the Parisi functional realized in terms of the Ruelle cascade, with parameters determined by $x_n(q)$, by $\mathcal{P}(x_n(q))$ (see Theorem 2.4.2 above). If N_k is a subsequence that realizes the $\liminf_{N\to\infty} A(x_N)$, then the proof is completed with the following lemma:

Lemma 2.4.7. *(eq. (3.84) [18])*

$$\liminf_{N\to\infty} A(x_N) = \lim_{k\to\infty} A(x_{N_k}) = \lim_{n\to\infty}\mathcal{P}(x_n(q)) \geq \inf_x \mathcal{P}(x) \quad (2.105)$$

Chapter 3

Deterministic multi-species models

In this chapter we deal with deterministic multi-species models. A first instance of these models, with only two populations was introduced by Contucci and Gallo in [13], then also studied in [12] and [6].

We begin with a multi-layer deterministic Curie-Weiss model that is again an example of a multi-species model with a specific choice of the structure of the interactions. This model turns out to have a behaviour similar to that of an anti-ferromagnetic model. This is why Lemma 1.4.8 can be useful even in this case, in fact it provides the lower bound needed to find the pressure per particle in the thermodynamic limit.

After that, we introduce a more general theory in which the structure of the reduced interaction matrix between the various species is not specified. The only hypothesis needed are the positivity or negativity of the previously mentioned matrix. We proceed with the proof of the existence of the limit in these two cases, then we find it explicitly.

Finally, we conclude with the computation of the finite size corrections with the large deviations method already employed before, for a positive definite reduced interaction matrix. The modification induced by a change in the normalization of the interaction term are also found.

3.1 The multi-layer Curie-Weiss model

Consider a spin system of N particles divided in K layers $L_1, \dots, L_K$ of sizes $N_1, \dots, N_K$ respectively, such that $\sum_{p=1}^{K} N_p = N$. The spins in the layer L_p interact with all the spins in the layers L_{p-1}, L_{p+1} and only with them. We assume that the relative sizes of the layers are fixed to $N_p/N = \alpha_p \in (0,1)$ for every $p = 1, \dots, K$ as N goes to infinity.

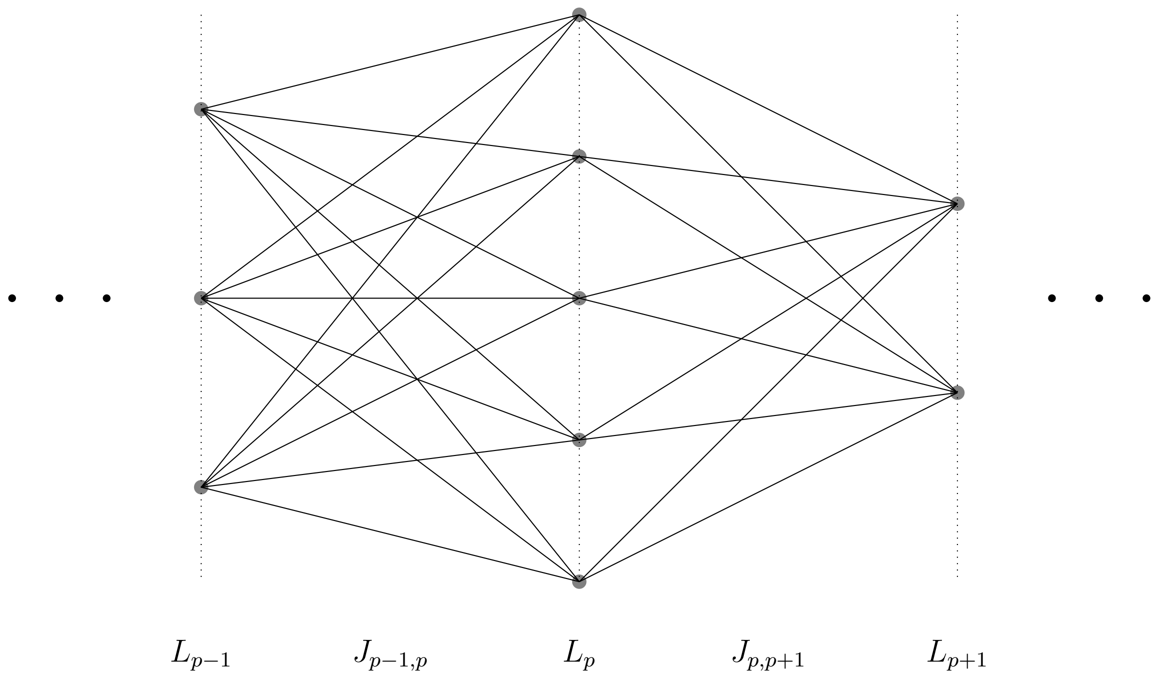


Figure 3.1: Scheme of the interactions between some layers.

The hamiltonian of the model is defined as follows:

Definition 3.1.1 (Multi-layer CW hamiltonian)**.** Given a set of couplings $J_{p,p+1} > 0$ and of magnetic fields $h_p \geq 0$, with $p = 1, \dots, K$, the hamiltonian of the multi-layer CW model is:

$$H_N^{ml}(\sigma; J, h) = -\frac{1}{N} \sum_{p=1}^{K-1} J_{p,p+1} \sum_{(i.j) \in L_p \times L_{p+1}} \sigma_i \sigma_j - \sum_{p=1}^{K} h_p \sum_{i \in L_p} \sigma_i \tag{3.1}$$

Remark 3.1.1. By adopting the following notation:

$$m_p = \frac{1}{N_p} \sum_{i \in L_p} \sigma_i \tag{3.2}$$

$$J'_{p,p+1} = \alpha_p J_{p,p+1} \alpha_{p+1} \tag{3.3}$$

the hamiltonian can be written in a simpler form:

$$H_N^{ml}(\sigma; J, h) = -N \sum_{p=1}^{K-1} J'_{p,p+1} m_p m_{p+1} - N \sum_{p=1}^{K} \alpha_p h_p m_p \tag{3.4}$$

Since the interaction matrix between the magnetizations has only non negative entries, Perron-Frobenius theorem guarantees that the dominant eigenvalue is strictly positive. However, the only non zero elements are off-diagonal, thus the model may still have a sort of antiferromagnetic behaviour. This fact leads us to think that we can find a lower bound by means of the convexity of an appropriate interpolating pressure, as already shown in the case of the anti-ferromagnetic CW model. Actually this is exactly the strategy used in the following theorem.

Theorem 3.1.1. *Let $r_1, \dots, r_{K-1}$ be a collection of positive numbers, and:*

$$\tilde{J}_p(r) = \frac{\alpha_{p-1}}{r_{p-1}} J_{p-1,p} + r_p J_{p,p+1} \alpha_{p+1} \tag{3.5}$$

with $J_{0,1} = J_{K,K+1} = 0$. Then, the pressure of the ferromagnetic multi-layer CW model in the thermodynamic limit is:

$$p^{ml}(J,h) = \lim_{N\to\infty} \frac{1}{N} \log \sum_{\sigma\in\Sigma_N} e^{-H_N^{ml}} = \inf_{r_1,\dots,r_{K-1}} \sum_{p=1}^{K} \alpha_p p^{CW}(\tilde{J}_p(r), h_p) \tag{3.6}$$

Proof. For the sake of clarity we divide the proof in two parts.

Part 1: Upper bound The upper bound can be easily found exploiting the positivity of the square:

$$\left(\sqrt{r_p} m_p - \frac{m_{p+1}}{\sqrt{r_p}}\right) \geq 0 \quad \Rightarrow \quad m_p m_{p+1} \leq \frac{1}{2}\left(r_p m_p^2 + \frac{m_{p+1}^2}{r_p}\right) \tag{3.7}$$

where we have introduced a set of positive numbers $r_1, \dots, r_{K-1}$. In this way we get immediately:

$$\sum_{p=1}^{K-1} J'_{p,p+1} m_p m_{p+1} \leq \sum_{p=1}^{K-1} \frac{J'_{p,p+1}}{2}\left(r_p m_p^2 + \frac{m_{p+1}^2}{r_p}\right) = \sum_{p=1}^{K} \alpha_p \frac{\tilde{J}_p(r)}{2} m_p^2 \tag{3.8}$$

To prove the last equality it is sufficient to shift the sum variable p and use the definitions of $\tilde{J}_p(r)$ and $J'_{p,p+1}$. Finally:

$$\begin{aligned} p_N^{ml} &= \frac{1}{N} \log \sum_{\sigma\in\Sigma_N} \exp\left[N \sum_{p=1}^{K-1} J'_{p,p+1} m_p m_{p+1} + N \sum_{p=1}^{K} \alpha_p h_p m_p\right] \leq \\ &\leq \frac{1}{N} \log \sum_{\sigma\in\Sigma_N} \exp\left[\sum_{p=1}^{K} N_p \frac{\tilde{J}_p(r)}{2} m_p^2 + \sum_{p=1}^{K} N_p h_p m_p\right] = \sum_{p=1}^{K} \alpha_p p_N^{CW}(\tilde{J}_p(r), h_p) \end{aligned} \tag{3.9}$$

Notice that the inequality is not uniform in N. On the contrary, as we shall see later, the lower bound is independent of N. After taking the lim sup and optimizing in the parameters r we get:

$$\limsup_{N\to\infty} p_N^{ml} \leq \inf_{r_1,\dots,r_{K-1}} \sum_{p=1}^{K} \alpha_p p^{CW}(\tilde{J}_p(r), h_p) \tag{3.10}$$

Part 2: Lower Bound In order to find a lower bound we use the interpolation technique, as suggested before. Let us define:

$$H_N^{free} = -\sum_{p=1}^{K} N_p m_p(\tilde{J}(r)_p x_p + h_p) \tag{3.11}$$

$$H_N(t) = H_N^{ml} t + H_N^{free}(1-t) \tag{3.12}$$

$$p_N(t) = \frac{1}{N} \log \sum_{\sigma\in\Sigma_N} e^{-H_N(t)} \tag{3.13}$$

$$p_N(0) = \sum_{p=1}^{K} \alpha_p \log 2\cosh(\tilde{J}(r)_p x_p + h_p) \tag{3.14}$$

$$p_N(1) = p_N^{ml} \tag{3.15}$$

Now, a simple application of the fundamental theorem of the integral calculus yields:

$$p_N^{ml} = p_N(0) + \int_0^1 dt\, \frac{dp_N(t)}{dt} = \sum_{p=1}^{K} \alpha_p \log 2\cosh(\tilde{J}(r)_p x_p + h_p) - \tag{3.16}$$

$$-\frac{1}{N}\int_0^1 \omega_{N,t}\left[\frac{dH_N(t)}{dt}\right] dt \tag{3.17}$$

Let us focus on the expectation in the integral:

$$\begin{aligned}
-\frac{1}{N}\omega_{N,t}\left[\frac{dH_N(t)}{dt}\right] &= \omega_{N,t}\left[-h_N^{CW} + h_N^{free}\right] = \\
&- \sum_{p=1}^{K} \alpha_p \tilde{J}_p(r) x_p \omega_{N,t}(m_p) + \sum_{p=1}^{K-1} J'_{p,p+1}\omega_{N,t}(m_p m_{p+1}) = \\
&= \sum_{p=1}^{K-1} J'_{p,p+1}\left[\omega_{N,t}(m_p m_{p+1}) - \frac{x_{p+1}}{r_p}\omega_{N,t}(m_{p+1}) - x_p r_p \omega_{N,t}(m_p)\right]
\end{aligned} \tag{3.18}$$

The last equality follows again from the definitions of J and $\tilde{J}$.

Now we choose the free parameters according to the following prescriptions:

$$\bar{x}_p = \tanh(\tilde{J}_p(r)\bar{x}_p + h_p) \tag{3.19}$$

$$\bar{x}_p^2 = \frac{\bar{x}_{p+1}^2}{\bar{r}_p^2} \tag{3.20}$$

and apply Lemma 1.4.8:

$$\begin{aligned} -\frac{1}{N}\omega_{N,t}\left[\frac{dH_N(t)}{dt}\right] \geq \sum_{p=1}^{K-1} J'_{p,p+1}\left[\bar{x}_p\bar{x}_{p+1} - \frac{\bar{x}_{p+1}^2}{\bar{r}_p} - \bar{x}_p^2\bar{r}_p\right] = -\sum_{p=1}^{K-1} J'_{p,p+1}\frac{\bar{x}_{p+1}^2}{\bar{r}_p} = \\ = -\sum_{p=1}^{K} \frac{J'_{p,p+1}}{2}\left(\frac{\bar{x}_{p+1}^2}{\bar{r}_p} + \bar{x}_p^2\bar{r}_p\right) = -\sum_{p=1}^{K-1}\frac{\alpha_p}{2}\left(\frac{\alpha_{p-1}}{\bar{r}_{p-1}}J_{p-1,p} + \bar{r}_p J_{p,p+1}\alpha_{p+1}\right)\bar{x}_p^2 = \\ = -\sum_{p=1}^{K}\alpha_p\frac{\tilde{J}_p(\bar{r})}{2}\bar{x}_p^2 \end{aligned} \tag{3.21}$$

Finally:

$$p_N^{ml} \geq \sum_{p=1}^{K}\alpha_p\left(\log 2\cosh(\tilde{J}(\bar{r})_p\bar{x}_p + h_p) - \frac{\tilde{J}_p(\bar{r})\bar{x}_p^2}{2}\right) = \sum_{p=1}^{K}\alpha_p p^{CW}(\tilde{J}_p(\bar{r}), h_p) \tag{3.22}$$

It turns out that the choice of $r_p = \bar{r}_p$ realizes the infimum of the previous expression, *i.e.*:

$$\psi(r) = \sum_{p=1}^{K}\alpha_p p^{CW}(\tilde{J}_p(r), h_p) \tag{3.23}$$

$$\frac{\partial\psi}{\partial r_p}(\bar{r}) = 0 \tag{3.24}$$

In fact, an explicit computation of the derivative yields:

$$\frac{\partial\psi}{\partial r_p}(\bar{r}) = \sum_{q=1}^{K}\alpha_q\left.\frac{\partial}{\partial r_p}\right|_{r=\bar{r}} p^{CW}(\tilde{J}_q(r), h_q) = \frac{1}{2}J'_{p,p+1}\left(\bar{x}_p^2 - \frac{\bar{x}_{p+1}^2}{\bar{r}_p^2}\right) \tag{3.25}$$

$\bar{r}$ exists because ψ is a convex function on $(0,\infty)^{K-1}$, indeed the pressure $p^{CW}(J,h)$ is convex and non-decreasing with respect to J and each coupling $\tilde{J}_p(r)$ is convex with respect to r. Moreover, each $\bar{r}_p$ is finite and non-zero, indeed $\psi(r)$ diverges both for $r_p \to 0$ and $r_p \to \infty$.

The lower bound can be written in this form:

$$p_N^{ml} \geq \inf_{r_1,\dots,r_{K-1}} \sum_{p=1}^{K} \alpha_p p^{CW}(\tilde{J}_p(r), h_p) \tag{3.26}$$

Notice that, as said before, it is uniform in N. This completes the proof.

□

3.2 Multi-species models

The systems discussed above are only a particular instance of *multi-species models*. In these type of models the population of spins is divided in, say, K subsets. These subsets are commonly called species, and we will denote them with S_p, where p labels the species. A spin of the p-th species interacts with coupling $J_{p,l}$ with another spin of the l-th species. We implicitly include interactions among spins of the same species, *i.e.* when $p = l$. We can conclude that the symmetric coupling matrix J_{ij} can be divided in blocks whose dimension is related to the dimension of the species.

$$J_{ij} = \begin{pmatrix} J_{11} & J_{12} & \cdots & J_{1K} \\ J_{21} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ J_{K1} & \cdots & \cdots & J_{KK} \end{pmatrix} \quad \text{with } i,j = 1,\dots,N \tag{3.27}$$

For our applications we will consider only positive diagonal elements or a definite reduced interaction matrix (as specified later). Furthermore, each species has its own local external magnetic field h_p, that is constant inside the species itself. At this point, we can formally rearrange magnetic fields in a vector:

$$h_i = \begin{pmatrix} h_1 \\ \vdots \\ h_K \end{pmatrix} \equiv \mathbf{h} \tag{3.28}$$

It is also convenient to define the following *local magnetizations* and to arrange them in a vector too:

$$m_p = \frac{1}{|S_p|} \sum_{i \in S_p} \sigma_i \quad \text{and} \quad \mathbf{m} = \begin{pmatrix} m_1 \\ \vdots \\ m_K \end{pmatrix} \tag{3.29}$$

Finally the ratio between the size $|S_p|$ of the species, and the total number of spins N, which will be denoted by α_p, is fixed. Hence, the thermodynamic limit will be performed maintaining those ratios. We are now ready for a formal definition.

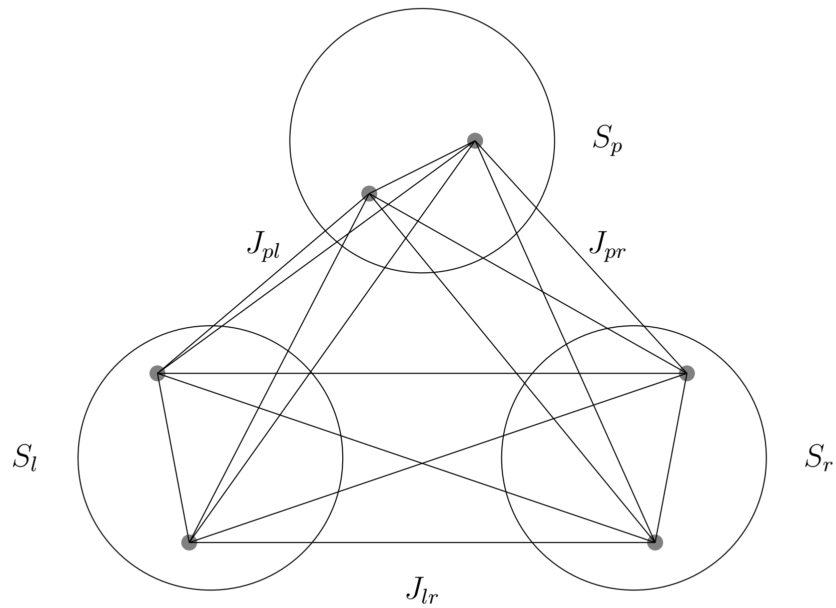


Figure 3.2: Scheme of the interactions with three species.

Definition 3.2.1. The hamiltonian of a K-species Curie-Weiss model is:

$$H_N^{ms}(\sigma) = -\frac{1}{2N}\sum_{i,j=1}^{N} J_{ij}\sigma_i\sigma_j - \sum_{i=1}^{N} h_i\sigma_i = -\frac{N}{2}\sum_{p,l=1}^{K}\alpha_p J_{pl}\alpha_l m_p m_l - N\sum_{p=1}^{K}\alpha_p h_p m_p \tag{3.30}$$

Remark 3.2.1. With little effort the hamiltonian can be rewritten with matrices and vectors. This allows us to get rid of indices for a great part of our work.

$$H_N(\sigma) = -\frac{N}{2}\langle \mathbf{m}, \Delta\mathbf{m}\rangle - N\langle \tilde{\mathbf{h}}, \mathbf{m}\rangle \tag{3.31}$$

where:

$$\Delta_{pl} = \alpha_p J_{pl}\alpha_l \quad \text{and} \quad \tilde{h}_p = \alpha_p h_p \tag{3.32}$$

Remark 3.2.2. In the multi-layer case the interaction matrix is of this form:

$$\begin{pmatrix} 0 & J_{12} & 0 & \cdots & \cdots & 0 \\ J_{21} & 0 & J_{23} & \ddots & \ddots & \vdots \\ 0 & J_{32} & 0 & J_{43} & \ddots & \vdots \\ \vdots & \ddots & J_{34} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 \end{pmatrix} \tag{3.33}$$

Only the elements immediately above and below the diagonal are non-zero.

3.2.1 Existence of the thermodynamic limit

The strategy to prove the existence of the thermodynamic limit is the very same used for the Curie-Weiss model. There are actually several methods, here we choose again the interpolation technique. It allows us to prove sub-additivity of the pressure. However, we still need to prove the boundedness of the pressure, in order to ensure that the limit is finite.

Lemma 3.2.1. *The sequence $p_N^{ms} = P_N^{ms}/N$ is bounded. More precisely:*

$$\log 2 + \frac{\inf_{i=1,\dots,N} J_{ii}}{2} \leq \frac{P_N^{ms}}{N} \leq \log 2 + \frac{K\|\Delta\|}{2} + \|\tilde{\mathbf{h}}\|\sqrt{K} \tag{3.34}$$

Proof. A repeated use of Schwarz's inequality, together with the definition of the norm of matrix Δ yields:

$$\begin{aligned} P_N^{ms} = \log \sum_{\sigma\in\Sigma_N} \exp\left(\frac{N}{2}\langle \mathbf{m}, \Delta\mathbf{m}\rangle + N\langle \tilde{\mathbf{h}}, \mathbf{m}\rangle\right) \leq \\ \leq \log \sum_{\sigma\in\Sigma_N} \exp\left(\frac{N}{2}\|\mathbf{m}\|^2\|\Delta\| + N\|\tilde{\mathbf{h}}\|\|\mathbf{m}\|\right) \leq \\ \leq \log \sum_{\sigma\in\Sigma_N} \exp\left(\frac{N}{2}K\|\Delta\| + N\|\tilde{\mathbf{h}}\|\sqrt{K}\right) = N\left(\log 2 + \frac{K}{2}\|\Delta\| + \|\tilde{\mathbf{h}}\|\sqrt{K}\right) \end{aligned} \tag{3.35}$$

Jensen's inequality immediately gives us the lower bound:

$$\begin{aligned} P_N^{ms} &= \log 2^N \omega_0 \left[\exp\left(\frac{1}{2N} \sum_{i,j=1}^N J_{ij}\sigma_i\sigma_j + \sum_{i=1}^N h_i\sigma_i \right) \right] \geq \\ &\geq \log 2^N \exp\left(\frac{1}{2N} \sum_{i,j=1}^N J_{ij}\omega_0\left(\sigma_i\sigma_j\right) \sum_{i=1}^N + h_i\omega_0(\sigma_i) \right) = N \left(\log 2 + \frac{\sum_i J_{ii}}{2N} \right) \end{aligned} \tag{3.36}$$

It is now sufficient to take the infimum with respect to the index i in the last sum. Dividing P_N^{ms} and both bounds by N the result is achieved. □

Remark 3.2.3. Now the idea is to interpolate our system with two others K-species systems. We need to pay attention to the fact that the ratios α's are fixed both for the original N particles system and for the new M_1 and M_2 particles systems, because, intuitively speaking, the interpolated systems must have the same characteristics, except for the number of particles. With this observation we are ready to prove the following theorem.

Theorem 3.2.2. *The sequence p_N^{ms} of the pressure per particle converges to its $\inf_N$, in the thermodynamic limit and the limit is finite:*

$$\lim_{N\to\infty} \frac{P_N^{ms}}{N} = \lim_{N\to\infty} p_N^{ms} = \inf_N p_N^{ms} \tag{3.37}$$

Proof. Taking into account the prescriptions listed in the previous remark, let us define the following interpolating hamiltonian and pressure:

$$H_N(t) = tH_N^{ms} + (1-t)[H_{M_1} + H_{M_2}] \quad \text{with } M_1 + M_2 = N \tag{3.38}$$

$$P_N(t) = \log \sum_{\sigma\in\Sigma_N} \exp\left(-H_N(t)\right) = \log \sum_{\sigma\in\Sigma_N} \exp\left(-tH_N^{ms} - (1-t)(H_{M_1} + H_{M_2})\right) \tag{3.39}$$

$$P_N(0) = P_{M_1} + P_{M_2} \tag{3.40}$$

$$P_N(1) = P_N^{ms} \tag{3.41}$$

Observe that, for $t = 1$, the spins inside the same species enjoy a permutation symmetry. They can be considered as indistinguishable when computing expectations and so on. Furthermore, the interpolating pressure is evidently convex in t.

Now we compute the first derivative of the interpolating pressure:

$$P_N'(t) = \omega_{N,t}\left[H^{ms}_{M_1} + H^{ms}_{M_2} - H^{ms}_N\right] = \omega_{N,t}\Bigg[-\frac{M_1}{2}\sum_{p,l=1}^{K}\alpha_p J_{pl}\alpha_l m_p^{(1)} m_l^{(1)} - \\ -\frac{M_2}{2}\sum_{p,l=1}^{K}\alpha_p J_{pl}\alpha_l m_p^{(2)} m_l^{(2)} + \frac{N}{2}\sum_{p,l=1}^{K}\alpha_p J_{pl}\alpha_l m_p m_l\Bigg] \leq P_N'(1) \quad (3.42)$$

The last inequality is obtained thanks to the convexity of the pressure. The notation $m_p^{(1)}$ has been employed to specify that the local magnetization has been computed only with the spins of the first partition. The same holds for the second partition. Now we separate the sum over the species in diagonal and off-diagonal terms. The diagonal terms are:

$$\frac{1}{2}\sum_{p=1}^{K} J_{pp}\alpha_p^2\omega_{N,1}\left[-M_1 m_p^{(1)\,2} - M_2 m_p^{(2)\,2} + N m_p^2\right] \quad (3.43)$$

Thanks to the positivity of the diagonal interaction matrix elements, the convexity of the square and:

$$N m_p = M_1 m_p^{(1)} + M_2 m_p^{(2)} \quad (3.44)$$

we can safely state that they are negative.

As far as the off-diagonal terms are concerned, it can be proved that they cancel each other, thanks to the permutation symmetry inside the species:

$$\frac{1}{2}\sum_{p\neq l} J_{pl}\Bigg[-\frac{1}{M_1}\sum_{(i,j)\in S_p^{(1)}\times S_l^{(1)}}\omega_{N,1}(\sigma_i\sigma_j) - \frac{1}{M_2}\sum_{(i,j)\in S_p^{(2)}\times S_l^{(2)}}\omega_{N,1}(\sigma_i\sigma_j) + \\ \frac{1}{N}\sum_{(i,j)\in S_p\times S_l}\omega_{N,1}(\sigma_i\sigma_j)\Bigg] = \frac{1}{2}\sum_{p\neq l} J_{pl}\left[-M_1\,\alpha_p\,\alpha_l\,\omega_{N,1}(\sigma_{1,p}\sigma_{1,l}) - \right. \\ \left. -M_2\,\alpha_p\,\alpha_l\,\omega_{N,1}(\sigma_{1,p}\sigma_{1,l}) + N\alpha_p\,\alpha_l\,\omega_{N,1}(\sigma_{1,p}\sigma_{1,l})\right] = 0 \quad (3.45)$$

Again with the notation $S_p^{(1)}$ we mean the set of spins of the p-th species in the first partition. Hence, we have proved that:

$$P_N'(t) \leq P_N'(1) \leq 0 \quad \Rightarrow \quad P_N(1) \leq P_N(0) \quad (3.46)$$

i.e. the pressure is sub-additive. Finally, Fekete's Lemma 1.1.1 implies the existence of the limit and the boundedness of the pressure per particle guarantees the finiteness of it.

□

A similar result can be proved exploiting the positivity, or negativity of the reduced interaction matrix Δ. For the rest of this chapter we will focus on these cases.

Proposition 3.2.3. *Consider the hamiltonian* (3.31). *If the reduced itneraction matrix Δ is (positive or negative) definite, the thermodynamic limit of the pressure exists.*

Proof. Let us examine the derivative of the interpolating pressure in (3.42), with particular attention to the square bracket. Assume the same definitions and notations in the proof of Theorem 3.2.2. We can rewrite $P_N'(t)$ using Δ:

$$P_N'(t) = \frac{N}{2}\omega_{N,t}\left[-\frac{M_1}{N}\langle \mathbf{m}^{(1)}, \Delta\mathbf{m}^{(1)}\rangle - \frac{M_2}{N}\langle \mathbf{m}^{(2)}, \Delta\mathbf{m}^{(2)}\rangle + \langle \mathbf{m}, \Delta\mathbf{m}\rangle\right] \tag{3.47}$$

If $\Delta > 0$ the quadratic forms above are convex in the set of local magnetizations. Thanks to the fact that the relative ratios of the species sizes are kept fixed, (3.44) holds again. This yields $P_N'(t) \leq 0$.

On the contrary if $\Delta < 0$ the very same quadratic forms are concave. This only reverts the previous inequality, $P_N'(t)$, but Fekete's Lemma 1.1.1 and the boundedness of P_N still imply the result. □

Remark 3.2.4. In our analysis we always dealt with Δ instead of the original $\mathbb{J}$ coupling matrix, where $\Delta_{pl} = J_{pl}\alpha_p\alpha_l$. The results obtained up to now do not change if we deal with $\mathbb{J}$. In fact, in this case, the ratios α_p can be arranged in a matrix $\hat{\alpha} = diag(\alpha_1, \ldots, \alpha_K)$ and the scalar products that always appear in the quadratic forms become:

$$\langle \mathbf{m}, \Delta\mathbf{m}\rangle = \langle \hat{\alpha}\mathbf{m}, \mathbb{J}(\hat{\alpha}\mathbf{m})\rangle \tag{3.48}$$

whose sign is established by the sign of $\mathbb{J}$, this time.

3.2.2 Solution of the model

The following lemmas will be useful to find the thermodynamic limit. The idea is to obtain one or both bounds through an appropriate interpolating hamiltonian. As we have already shown for multi-layer systems, the original hamiltonian is interpolated with a completely separable hamiltonian, related to a free system whose magnetic field is accurately chosen.

The variational expression obtained in the calculations, contrary to what happens for one species systems, is not so simple to deal with, because the equilibrium values of the parameters of the expression turn out to be dependent on each other.

Lemma 3.2.4. *The pressure per particle induced by the hamiltonian* (3.31) *can be written with the following sum rule:*

$$p_N^{ms}(\Delta, \mathbf{h}) = \sum_{p=1}^{K} \alpha_p \log \left[2\cosh \left(\sum_{l=1}^{K} J_{pl}\alpha_l x_l + h_p \right) \right] - \frac{1}{2}\langle \mathbf{x}, \Delta\mathbf{x} \rangle + R_N(\Delta; \mathbf{x}) \quad (3.49)$$

$$R_N(\Delta; \mathbf{x}) = \frac{1}{2}\int_0^1 dt\, \omega_{N,t}\left[\langle \mathbf{m} - \mathbf{x}, \Delta(\mathbf{m} - \mathbf{x})\rangle\right] \quad (3.50)$$

Proof. Let us define the hamiltonian:

$$H_N^{free}(\sigma; \Delta, \mathbf{h}, \mathbf{x}) = -N\langle \Delta\mathbf{x} + \tilde{\mathbf{h}}, \mathbf{m}\rangle \quad (3.51)$$

which clearly describes a free system. Now the interpolating hamiltonian that yields the result is:

$$H_N(t) = tH_N^{ms} + (1-t)H_N^{free} \quad (3.52)$$

$$(3.53)$$

Its pressure $p_N(t)$ is such that $p_N(1) = p_N^{ms}$ and:

$$\begin{aligned} p_N(0) = \frac{1}{N}\log \sum_{\sigma\in\Sigma_N} e^{N\langle \Delta\mathbf{x} + \tilde{\mathbf{h}}, \mathbf{m}\rangle} &= \\ = \frac{1}{N}\log \sum_{\sigma\in\Sigma_N} \exp\left[\sum_{p=1}^{K} N_p \left(\sum_{l=1}^{K} J_{pl}\alpha_l x_l + h_p \right) m_p \right] &= \\ = \frac{1}{N}\sum_{p=1}^{K} \log \sum_{\sigma\in\Sigma_{N_p}} \exp\left[N_p \left(\sum_{l=1}^{K} J_{pl}\alpha_l x_l + h_p \right) m_p \right] &= \\ = \sum_{p=1}^{K} \alpha_p \log \left[2\cosh\left(\sum_{l=1}^{K} J_{pl}\alpha_l x_l + h_p \right) \right] & \end{aligned} \quad (3.54)$$

Furthermore:

$$p_N'(t) = -\frac{1}{N}\omega_{N,t}\left[H_N^{ms} - H_N^{free}\right] = \frac{1}{2}\omega_{N,t}\left[\langle \mathbf{m}, \Delta\mathbf{m}\rangle - 2\langle \mathbf{m}, \Delta\mathbf{x}\rangle\right] =$$
$$= -\frac{1}{2}\langle \mathbf{x}, \Delta\mathbf{x}\rangle + \frac{1}{2}\omega_{N,t}\left[\langle \mathbf{m}-\mathbf{x}, \Delta(\mathbf{m}-\mathbf{x})\rangle\right] \quad (3.55)$$

Now, using the theorem of integral calculus we get:

$$p_N^{ms} = p_N(0) - \frac{1}{2}\langle \mathbf{x}, \Delta\mathbf{x}\rangle + \frac{1}{2}\int_0^1 dt\, \omega_{N,t}\left[\langle \mathbf{m}-\mathbf{x}, \Delta(\mathbf{m}-\mathbf{x})\rangle\right] \quad (3.56)$$

Finally, inserting $p_N(0)$ we get the result. □

Lemma 3.2.5. *Consider the variational function:*

$$p_{var}(\Delta, \mathbf{h}; \mathbf{x}) = -\frac{1}{2}\langle \mathbf{x}, \Delta\mathbf{x}\rangle + \sum_{p=1}^{K}\alpha_p \log\left[2\cosh\left(\sum_{l=1}^{K} J_{pl}\alpha_l x_l + h_p\right)\right] \quad (3.57)$$

with fixed Δ and $\mathbf{h}$, then:

- *if $\Delta > 0$, $p_{var}(\Delta, \mathbf{h}; \mathbf{x})$ is bounded only from above;*
- *if $\Delta < 0$, $p_{var}(\Delta, \mathbf{h}; \mathbf{x})$ is bounded only from below.*

Proof. Suppose $\Delta > 0$. Furthermore, without loss of generality, we can assume for the moment that $\sum_{l=1}^{K} J_{pl}\alpha_l x_l + h_p \geq 0$. Then:

$$\log\left[2\cosh\left(\sum_{l=1}^{K} J_{pl}\alpha_l x_l + h_p\right)\right] = \log\left[\exp\left(\sum_{l=1}^{K} J_{pl}\alpha_l x_l + h_p\right) + \right.$$
$$\left.\exp\left(-\sum_{l=1}^{K} J_{pl}\alpha_l x_l - h_p\right)\right] \leq \sum_{l=1}^{K} J_{pl}\alpha_l x_l + h_p + \log 2 \quad (3.58)$$

Hence:

$$p_{var}(\Delta, \mathbf{h}; \mathbf{x}) \leq -\frac{1}{2}\langle \mathbf{x}, \Delta\mathbf{x}\rangle + \sum_{p=1}^{K}\alpha_p\left(\sum_{l=1}^{K} J_{pl}\alpha_l x_l + h_p\right) + K\log 2 \quad (3.59)$$

The r.h.s. is a quadratic concave function, hence it is clearly bounded from above. Moreover, one can easily see that the limit for large $\|\mathbf{x}\|$ yields $-\infty$. Consequently, p_{var} inherits the same properties.

Let us now discuss the case $\Delta < 0$. The quadratic form $-\langle \mathbf{x}, \Delta\mathbf{x}\rangle$ is always positive. Thus $p_{var} \geq 0$. Finally, both terms in (3.57) approach $+\infty$ when $\|\mathbf{x}\| \to +\infty$. □

We are now ready to state the theorem that contains the solution of the model in the two exmined cases.

Theorem 3.2.6 (Solution to the Multi-species model)**.** *The limit of the pressure induced by the hamiltonian* (3.31) *is:*

- *if* $\Delta > 0$, $\lim_{N\to\infty} p_N^{ms}(\Delta, \mathbf{h}) = \sup_{\mathbf{x}} p_{var}(\Delta, \mathbf{h}; \mathbf{x})$*;*
- *if* $\Delta < 0$, $\lim_{N\to\infty} p_N^{ms}(\Delta, \mathbf{h}) = \inf_{\mathbf{x}} p_{var}(\Delta, \mathbf{h}; \mathbf{x})$.

Proof. We divide the proof for the two cases.

Positive definite interaction matrix: $\Delta > 0$ The lower bound can be immediately found thanks to Lemma 3.2.4. In fact, remembering that $\langle \mathbf{x}, \Delta\mathbf{x}\rangle \geq 0$, the rest $R_N(\Delta; \mathbf{x})$ has the same sign because of the positivity of the expectation functional. Hence:

$$p_N^{ms}(\Delta, \mathbf{h}) \geq p_{var}(\Delta, \mathbf{h}; \mathbf{x}) \tag{3.60}$$

Optimizing with respect to $\mathbf{x}$:

$$\liminf_{N\to\infty} p_N^{ms}(\Delta, \mathbf{h}) \geq \sup_{\mathbf{x}} p_{var}(\Delta, \mathbf{h}; \mathbf{x}) \tag{3.61}$$

The upper bound is obtained as follows. Let us write down the partition function:

$$Z_N^{ms} = \sum_{\sigma\in\Sigma_N} \overbrace{\sum_{\mathbf{x}\in M_N} \delta_{\mathbf{x},\mathbf{m}}}^{1} \exp\left(\frac{N}{2}\langle \mathbf{m}, \Delta\mathbf{m}\rangle + N\langle \tilde{\mathbf{h}}, \mathbf{m}\rangle\right) \tag{3.62}$$

where M_N is the magnetization spectrum. Thanks to Kronecker's delta we can linearize the quadratic term:

$$\langle \mathbf{m} - \mathbf{x}, \Delta(\mathbf{m} - \mathbf{x})\rangle = 0 \quad \Rightarrow \quad \langle \mathbf{m}, \Delta\mathbf{m}\rangle = 2\langle \mathbf{m}, \Delta\mathbf{x}\rangle - \langle \mathbf{x}, \Delta\mathbf{x}\rangle \tag{3.63}$$

Inserting it into the partition function, and imposing $\delta_{\mathbf{x},\mathbf{m}} \leq 1$:

$$\begin{aligned} Z_N^{ms} &\leq \sum_{\sigma\in\Sigma_N}\sum_{\mathbf{x}\in M_N} \exp\left(-\frac{N}{2}\langle \mathbf{x}, \Delta\mathbf{x}\rangle + N\langle \Delta\mathbf{x} + \tilde{\mathbf{h}}, \mathbf{m}\rangle\right) = \\ &= \sum_{\mathbf{x}\in M_N} \exp\left(-\frac{N}{2}\langle \mathbf{x}, \Delta\mathbf{x}\rangle + N\sum_{p=1}^{K} \alpha_p \log 2\cosh\left(\sum_{l=1}^{K} J_{pl}\alpha_l x_l + h_p\right)\right) \leq \\ &\leq \left[\prod_{p=1}^{K} (N_p + 1) \exp\left(N\alpha_p \log 2\cosh\left(\sum_{l=1}^{K} J_{pl}\alpha_l x_l + h_p\right)\right)\right] \exp\left(-\frac{N}{2}\langle \mathbf{x}, \Delta\mathbf{x}\rangle\right) \end{aligned} \tag{3.64}$$

Taking the log and dividing by N we get:

$$p_N^{ms} \leq -\frac{1}{2}\langle \mathbf{x}, \Delta\mathbf{x}\rangle + \sum_{p=1}^{K}\left[\alpha_p \log 2\cosh\left(\sum_{l=1}^{K} J_{pl}\alpha_l x_l + h_p\right) + \frac{\log(N_p+1)}{N}\right] \leq$$

$$\leq \sup_{\mathbf{x}} p_{var}(\Delta, \mathbf{h}; \mathbf{x}) + \sum_{p=1}^{K} \frac{\log(N_p+1)}{N} \quad (3.65)$$

Hence:

$$\limsup_{N\to\infty} p_N^{ms} \leq \sup_{\mathbf{x}} p_{var}(\Delta, \mathbf{h}; \mathbf{x}) \quad (3.66)$$

The existence of these suprema is guaranteed by the previous Lemma.

Negative definite interaction matrix: $\Delta < 0$ Again, one bound, the upper one this time. The rest is now negative:

$$p_N^{ms}(\Delta, \mathbf{h}) \leq p_{var}(\Delta, \mathbf{h}; \mathbf{x}) \quad \Rightarrow \quad p_N^{ms}(\Delta, \mathbf{h}) \leq \inf_{\mathbf{x}} p_{var}(\Delta, \mathbf{h}; \mathbf{x}) \quad (3.67)$$

The lower bound is obtained thanks to the convexity of the interpolating pressure (see Lemma 1.4.8). This property tells us that:

$$R_N(\Delta; \mathbf{x}) \geq \frac{1}{2}\omega_{N,0}\left[\langle \mathbf{m}-\mathbf{x}, \Delta(\mathbf{m}-\mathbf{x})\rangle\right] = \frac{1}{2}\sum_{p,l=1}^{K} J_{pl}\alpha_p\alpha_l\omega_{N,0}\left[(m_p - x_p)(m_l - x_l)\right] \quad (3.68)$$

Let us now separate the diagonal and off-diagonal terms:

$$R_N(\Delta; \mathbf{x}) \geq \frac{1}{2}\sum_{p=1}^{K} J_{pp}\alpha_p^2\omega_{N,0}\left[(m_p - x_p)^2\right] + \quad (3.69)$$

$$+\sum_{p\neq l}^{K} J_{pl}\alpha_p\alpha_l\omega_{N,0}\left[(m_p - x_p)\right]\omega_{N,0}\left[(m_l - x_l)\right] \quad (3.70)$$

$\omega_{N,0}$ is the measure of a free system, it is thus possible to compute expectations explicitly. If we denote the first spin of the species l by $\sigma_{1,l}$:

$$\omega_{N,0}(m_l) = \frac{1}{\prod_{p=1}^{K}\left(2\cosh\left(\sum_{l=1}^{K} J_{pl}\alpha_l x_l + h_p\right)\right)^{N_p}} \sum_{\sigma\in\Sigma_N} \sigma_{1,l}\exp\left(N\langle \Delta\mathbf{x} + \tilde{\mathbf{h}}, \mathbf{m}\rangle\right) =$$

$$= \tanh\left(\sum_{l=1}^{K} J_{pl}\alpha_l x_l + h_p\right) \quad (3.71)$$

Now we choose as parameters $\mathbf{x}$ those minimizing $p_{var}(\Delta, \mathbf{h}; \mathbf{x})$. Deriving w.r.t. a generic x_p we get the consistency relation:

$$\bar{x}_p = \tanh\left(\sum_{l=1}^{K} J_{pl}\alpha_l\bar{x}_l + h_p\right) \tag{3.72}$$

Hence:

$$\omega_{N,0}(m_l) = \bar{x}_p \tag{3.73}$$

$$\omega_{N,0}[(m_p - \bar{x}_p)^2] = \bar{x}_p^2 - 2\bar{x}_p^2 + \frac{1}{N_p} + \frac{N_p - 1}{N_p}\bar{x}_p^2 = \frac{1 - \bar{x}_p^2}{N_p} \tag{3.74}$$

Gathering altogether:

$$p_N^{ms}(\Delta, \mathbf{h}) \geq p_{var}(\Delta, \mathbf{h}; \mathbf{x}) + \frac{1}{2N}\sum_{p=1}^{K} J_{pp}\alpha_p(1 - \bar{x}_p^2) \geq \tag{3.75}$$

$$\geq \inf_{\mathbf{x}} p_{var}(\Delta, \mathbf{h}; \mathbf{x}) + \frac{1}{2N}\sum_{p=1}^{K} J_{pp}\alpha_p(1 - \bar{x}_p^2) \tag{3.76}$$

This proves the claim

□

3.2.3 Finite size corrections

Lemma 3.2.7 (Generalized Hubbard-Stratonovič transform)**.** *Let Δ be a $K \times K$ symmetric, positive definite matrix, $\mathbf{m} \in \mathbb{R}^K$ and $a \in \mathbb{R}$. Then:*

$$\exp\left(\frac{a}{2}\langle \mathbf{m}, \Delta\mathbf{m}\rangle\right) = \int_{\mathbb{R}^K} d^K x \sqrt{\frac{a^K det\Delta}{(2\pi)^K}} \exp\left(-\frac{a}{2}\langle \mathbf{x}, \Delta\mathbf{x}\rangle + a\langle \mathbf{x}, \Delta\mathbf{m}\rangle\right) \tag{3.77}$$

Proof. The proof consists in the computation of the gaussian integral on the r.h.s. Observe that, since Δ is symmetric and positive definite it can always be written as the product of an appropriate matrix A with A^T. This allows us to recast the argument of the exponential:

$$-\frac{a}{2}\langle \mathbf{x}, \Delta\mathbf{x}\rangle + a\langle \mathbf{x}, \Delta\mathbf{m}\rangle = -\frac{a}{2}\langle A\mathbf{x}, A\mathbf{x}\rangle + a\langle A\mathbf{x}, A\mathbf{m}\rangle - \frac{a}{2}\langle A\mathbf{m}, A\mathbf{m}\rangle + \frac{a}{2}\langle A\mathbf{m}, A\mathbf{m}\rangle \tag{3.78}$$

We can thus change the integration variable: $A(\mathbf{x}-\mathbf{m})=\mathbf{y}$. The Jacobian of this transformation is precisely $\det A=\sqrt{det\Delta}$.

$$\int_{\mathbb{R}^K} d^K x\sqrt{\frac{a^K det\Delta}{(2\pi)^K}}\exp\left(-\frac{a}{2}\langle\mathbf{x},\Delta\mathbf{x}\rangle+a\langle\mathbf{x},\Delta\mathbf{m}\rangle\right)=$$

$$=\exp\left(\frac{a}{2}\langle\mathbf{m},\Delta\mathbf{m}\rangle\right)\underbrace{\int_{\mathbb{R}^K} d^K x\sqrt{\frac{a^K det\Delta}{(2\pi)^K}}\exp\left(-\frac{a}{2}\langle A(\mathbf{x}-\mathbf{m}),A(\mathbf{x}-\mathbf{m})\rangle\right)}_{1} \quad (3.79)$$

□

Lemma 3.2.8 (Generalized Laplace's asymptotic estimate)**.** *Let $F:\mathbb{A}\subseteq\mathbb{R}^K\longrightarrow\mathbb{R}$ be a twice differentiable function with a global maximum in x_0, and $g(x)$ an analytic function in a neighbourhood of x_0. The following estimate holds:*

$$\int_{A\subseteq\mathbb{R}^K} d^K x\, e^{NF(x)}g(x)=\sqrt{\frac{(2\pi)^K}{N^K\, det\,\partial^2 F(x_0)}}e^{NF(x_0)}\left(g(x_0)+O\left(\frac{1}{N}\right)\right) \quad (3.80)$$

With these two lemmas we are now ready to compute the first order finite size correction to the pressure.

Theorem 3.2.9. *Consider the case of a positive definite interaction matrix Δ. If the number of degenerate global maxima of $p_{var}(\Delta,\mathbf{h};\mathbf{x})$ M is finite, and the interaction matrix is positive definite, then:*

$$p_N(\Delta,\mathbf{h})=\sup_{\mathbf{x}} p_{var}(\Delta,\mathbf{h};\mathbf{x})-\frac{1}{2N}\log\left[\frac{|det\,\partial^2 p_{var}(\Delta,\mathbf{h};\mathbf{x}_0)|}{det\Delta\, M^2}\right]+O\left(\frac{1}{N^2}\right) \quad (3.81)$$

where $\partial^2 p_{var}$ is the Hessian matrix of p_{var} and $\mathbf{x}_0$ one of the degenerate global maxima.

Proof. We will use the generalized Hubbard-Stratonovič transform to linearize the quadratic interaction term. Let us write down the partition function.

$$Z_N=\sum_{\sigma\in\Sigma_N}\exp\left[\frac{N}{2}\langle\mathbf{m},\Delta\mathbf{m}\rangle+N\langle\tilde{\mathbf{h}},\mathbf{m}\rangle\right]=$$

$$=\sum_{\sigma\in\Sigma_N}\int_{\mathbb{R}^K} d^K x\sqrt{\frac{N^K det\Delta}{(2\pi)^K}}\exp\left[-\frac{N}{2}\langle\mathbf{x},\Delta\mathbf{x}\rangle+N\langle\Delta\mathbf{x}+\tilde{\mathbf{h}},\mathbf{m}\rangle\right] \quad (3.82)$$

Computing the sum, what remains at the exponent is nothing but N times the variational pressure:

$$Z_N = \int_{\mathbb{R}^K} d^K x \sqrt{\frac{N^K det\,\Delta}{(2\pi)^K}} \exp\left(N p_{var}(\Delta, \mathbf{h}; \mathbf{x})\right) \tag{3.83}$$

Thanks to Laplace's estimate we can write:

$$Z_N = M \sqrt{\frac{N^K det\,\Delta}{(2\pi)^K}} \exp\left(N p_{var}(\Delta, \mathbf{h}; \mathbf{x}_0)\right) \times \tag{3.84}$$

$$\times \sqrt{\frac{(2\pi)^K}{N^K |det\,\partial^2 p_{var}(\Delta, \mathbf{h}; \mathbf{x}_0)|}} \left(1 + O\left(\frac{1}{N}\right)\right) \tag{3.85}$$

where we have taken into account the possibility to have M degenerate global maxima of the variational pressure; $\mathbf{x}_0$ is simply one of them.

Finally, taking the log and dividing by N we get the result:

$$p_N(\Delta, \mathbf{h}) = p_{var}(\Delta, \mathbf{h}; \mathbf{x}_0) - \frac{1}{2N} \log\left[\frac{|det\,\partial^2 p_{var}(\Delta, \mathbf{h}; \mathbf{x}_0)|}{det\,\Delta\, M^2}\right] + O\left(\frac{1}{N^2}\right) \tag{3.86}$$

□

3.2.4 Normalization effects

Theorem 3.2.10 (Normalization effects on p_N^{ms})**.** *Consider the modified hamiltonian:*

$$H_{N,c}^{ms}(\sigma) = -\frac{1}{2(N+c)} \sum_{i,j=1}^{N} J_{ij}\sigma_i\sigma_j - \sum_{i=1}^{N} h_i\sigma_i = -\frac{N^2}{2(N+c)} \langle \mathbf{m}, \Delta\mathbf{m}\rangle - N\langle \tilde{\mathbf{h}}, \mathbf{m}\rangle \tag{3.87}$$

where $c \in \mathbb{R}$ and $\Delta > 0$. Then the finite size corrections to the pressure modify as follows:

$$p_{N,c}^{ms} = \frac{1}{N} \log Z_{N,c}^{ms} = p_{var}(\Delta, \mathbf{h}; \mathbf{z}_0) + \frac{C_N^{(1)}}{N} + O\left(\frac{1}{N^2}\right) \tag{3.88}$$

$$C_N^{(1)} = \log M + \frac{1}{2}\left[\log\left(\frac{det\,\Delta}{|det\,\partial^2 p_{var}(\Delta, \mathbf{h}; \mathbf{z}_0)|}\right) - c\langle \mathbf{z}_0, \Delta\mathbf{z}_0\rangle\right] \tag{3.89}$$

Hence if c is sufficiently large the correction becomes negative.

Proof. We proceed as in the proof of the finite size corrections. Let us start from the partition function and then use Hubbard-Stratonovič transform.

$$
\begin{aligned}
Z^{ms}_{N,c} = \sum_{\sigma\in\Sigma_N} \exp\left[\frac{N^2}{2(N+c)}\langle \mathbf{m}, \Delta\mathbf{m}\rangle + N\langle\tilde{\mathbf{h}},\mathbf{m}\rangle\right] = \\
= \sum_{\sigma\in\Sigma_N}\int_{\mathbb{R}^K} d^K x \sqrt{\frac{N^{2K} det\,\Delta}{(2\pi)^K (N+c)^K}} \times \\
\times \exp\left[-\frac{N^2}{2(N+c)}\langle \mathbf{x}, \Delta\mathbf{x}\rangle + \frac{N^2}{2(N+c)}\langle \Delta\mathbf{x}, \mathbf{m}\rangle + N\langle\tilde{\mathbf{h}},\mathbf{m}\rangle\right]
\end{aligned}
\tag{3.90}
$$

with the change $\mathbf{z} = N/(N+c)\mathbf{x}$ we get:

$$
\begin{aligned}
Z^{ms}_{N,c} = \sum_{\sigma\in\Sigma_N}\int_{\mathbb{R}^K} d^K z \sqrt{\frac{(N+c)^K\, det\,\Delta}{(2\pi)^K}} \exp\left[-\frac{(N+c)}{2}\langle \mathbf{z}, \Delta\mathbf{z}\rangle + N\langle\Delta\mathbf{z}+\tilde{\mathbf{h}},\mathbf{m}\rangle\right] = \\
= \int_{\mathbb{R}^K} d^K z \sqrt{\frac{(N+c)^K\, det\,\Delta}{(2\pi)^K}} \exp\left(Np_{var}(\Delta,\mathbf{h};\mathbf{z})\right) \exp\left(-\frac{c}{2}\langle\mathbf{z},\Delta\mathbf{z}\rangle\right) = \\
= M\sqrt{\frac{(N+c)^K\, det\,\Delta}{(2\pi)^K}} \left(Np_{var}(\Delta,\mathbf{h};\mathbf{z}_0)\right)\left[\exp\left(-\frac{c}{2}\langle\mathbf{z}_0,\Delta\mathbf{z}_0\rangle\right) + O\left(\frac{1}{N}\right)\right] \times \\
\times \sqrt{\frac{(2\pi)^K}{N^K |det\,\partial^2 p_{var}(\Delta,\mathbf{h};\mathbf{z}_0)|}}
\end{aligned}
\tag{3.91}
$$

where as usual M is the number of degenerate global maxima, and z_0 is one of those global maxima. Use has been made of Lemma 3.2.8. Thus taking the log divided by N one gets:

$$
\begin{aligned}
p^{ms}_{N,c} = p_{var}(\Delta,\mathbf{h};\mathbf{z}_0) + \frac{\log M}{N} + \frac{K}{2N}\log\left(\frac{N+c}{N}\right) + \\
+ \frac{1}{2N}\left[\log\left(\frac{det\,\Delta}{|det\,\partial^2 p_{var}(\Delta,\mathbf{h};\mathbf{z}_0)|}\right) - c\langle\mathbf{z}_0,\Delta\mathbf{z}_0\rangle\right] + O\left(\frac{1}{N^2}\right)
\end{aligned}
\tag{3.92}
$$

Hence the first correction modifies to:

$$
C^{(1)}_N = \log M + \frac{1}{2}\left[\log\left(\frac{det\,\Delta}{|det\,\partial^2 p_{var}(\Delta,\mathbf{h};\mathbf{z}_0)|}\right) - c\langle\mathbf{z}_0,\Delta\mathbf{z}_0\rangle\right] \tag{3.93}
$$

□

Corollary 3.2.11 (Stability for change of normalization). *p_N^{ms} and $p_{N,c}^{ms}$, as defined before, have the same thermodynamic limit:*

$$\lim_{N\to\infty} p_N^{ms} = \lim_{N\to\infty} p_{N,c}^{ms} \tag{3.94}$$

Proof. c is contained only in $C_N^{(1)}$, the result follows immediately from the previous theorem.

The same result could be obtained without the exact solution of the model and the finite size corrections via interpolation. □

Remark 3.2.5. Notice that we always relied on the fact that M is finite. This in particular can be explicitly proved for the two-species version of the model as done in [13], where it is shown that the number of maxima is bounded from above by 5.

Chapter 4

Multi-species disordered models

These type of models are quite recent and were introduced and studied by Barra, Contucci, Mingione and Tantari in [7] and Panchenko in [19]. The first authors found a replica symmetry breaking *ansatz* for the thermodynamic limit and proved that it bounds the pressure of the model from above. D. Panchenko proved the other bound, thus completing the solution of the model.

Up to now we are only able to deal with elliptic models. A multi-species disordered model is said to be elliptic if the matrix that we will denote by Δ^2, defined by the following:

$$\Delta^2_{rs}\delta_{rr'}\delta_{ss'}\delta_{ik}\delta_{jl} = \mathbb{E}[J^{rs}_{ij}J^{r's'}_{kl}] \quad \text{with } r,s = 1,\ldots,K \text{ and } \Delta^2_{rs} = \Delta^2_{sr} \tag{4.1}$$

is positive definite. J^{rs}_{ij} is the centered gaussian centered coupling between two spin σ_i, $i \in \Lambda_r$ and σ_j, $j \in \Lambda_s$, and Λ_r, Λ_s are the species, as defined below.

Contrary to what happens for deterministic multi-species models, Δ^2 cannot be negative definite. This is a consequence of the Perron-Frobenius theorem that states that a non-negative matrix has a dominant positive eigenvalue. Therefore, the only interesting models are elliptic or hyperbolic, when the matrix is not definite.

An important instance of hyperbolic model is the multi-layer SK model, that will be studied in Chapter 5. In hyperbolic models, one could proceed with an extension of the replica trick to have an idea of what the solution should be, keeping in mind that it is not a rigorous procedure. Unfortunately, the linearization of quadratic terms through Hubbard-Stratonovič trasnform in the exponent of the partition function relies on the positivity of the covariance matrix Δ^2. This makes this approach useless, as far as we know, for hyperbolic cases.

4.1 The model, elliptic case

Let us now begin with a more robust introduction to the elliptic multi-species disordered model. As for the deterministic version, consider a set of indices Λ, whose cardinality is $|\Lambda| = N$, equal to the number of spins included in the system. Imagine to divide it in K species, *i.e.* K disjoint sets Λ_r, $r = 1, \ldots, K$ such that $\uplus_{r=1}^{K} \Lambda_r = \Lambda$. This time the couplings are not *i.i.d.*, though they are still gaussian and centered. As mentioned in the introduction of this chapter $J_{ij}^{rs} \overset{\text{iid}}{\sim} \mathcal{N}(0, \Delta_{rs}^2)$ with $i \in \Lambda_r$, $j \in \Lambda_s$. We are now ready to write the hamiltonian.

Definition 4.1.1 (Disordered Mutli-species model hamiltonian)**.** Let $J_{ij}^{rs} \overset{\text{iid}}{\sim} \mathcal{N}(0, \Delta_{rs}^2)$. The hamiltonian of a disordered multi-species model is:

$$H_N^{dms}(\sigma; J) = -\frac{1}{\sqrt{N}} \sum_{r,s=1}^{K} \sum_{(i,j)\in\Lambda_r\times\Lambda_s} J_{ij}^{rs} \sigma_i \sigma_j \tag{4.2}$$

If Δ_{rs}^2 is positive definite, the model is *elliptic*.

Remark 4.1.1. The model could be equivalently defined by considering a family of gaussian r.v. H_σ^{dms} indexed by the spin configuration whose covariance is:

$$\mathbb{E}[H_\sigma^{dms} H_\tau^{dms}] = N \sum_{r,s=1}^{K} \alpha_r \Delta_{rs}^2 \alpha_s q_r(\sigma,\tau) q_s(\sigma,\tau) \equiv \langle \tilde{\mathbf{q}}(\sigma,\tau), \Delta^2 \tilde{\mathbf{q}}(\sigma,\tau) \rangle \tag{4.3}$$

$$q_r(\sigma,\tau) = \frac{1}{N_r} \sum_{i\in\Lambda_r} \sigma_i \tau_i \tag{4.4}$$

where we have introduced a vector notation, inspired by the deterministic case, with $\tilde{q}_s = \alpha_s q_s$. α's are nothing but the ratios $|\Lambda_s|/|\Lambda| = N_s/N$. It can be immediately verified by inspection that the family (4.2) has exactly the covariance (4.3). In the elliptic case the covariance is always positive, thanks to the positivity of Δ^2.

In the following, we will introduce also deterministic weights induced by the presence of an external magnetic field which we assume to be locally constant inside each species. Thus we recover again the vector notation, and write $\mathbf{h}$ when useful.

4.1.1 Existence of the thermodynamic limit

The proof of the existence of the limit follows the same steps, with slight modifications that take into account the presence of different species. Again, in the elliptic case, the super-additivity of the pressure is still guaranteed thanks to the convexity of the quadratic form defined through Δ^2.

Theorem 4.1.1. *The pressure of the disordered multi-species elliptic model:*

$$P_N^{dms}(\beta,\mathbf{h}) = \mathbb{E}\log\left[\sum_{\sigma\in\Sigma_N}\exp\left(-\beta H_N^{dms}(\sigma;J)\right)W_N(\sigma;\beta,\mathbf{h})\right] \tag{4.5}$$

$$W_N(\sigma;\beta,\mathbf{h}) = \prod_{r=1}^{K}W_r(\sigma;\beta,h_r) = \prod_{r=1}^{K}\exp\left(\beta h_r\sum_{i\in\Lambda_r}\sigma_i\right) \tag{4.6}$$

is super-additive. Hence the limit when $N\to\infty$ of the pressure per particle sequence P_N^{dms}/N exists.

Proof. The proof again can be performed through interpolation. Imagine to have a system of N particles partitioned in two subsystems N_1 and N_2, The three systems, the original one and its two partitions are characterized by the same ratios α between the species. Consider now the two independent gaussian families: $H_N^{dms}(\sigma;J)$ and $\tilde{H}_N(\sigma;\tilde{J}) = H_{N_1}^{dms}(\sigma;\tilde{J}) + H_{N_2}^{dms}(\sigma;\tilde{J})$. Since the two subsystems both must reproduce a smaller copy of the N particle system, the covariances of the couplings are untouched. Let us compute the following covariance:

$$\begin{aligned}\mathbb{E}[\tilde{H}_N(\sigma;\tilde{J})\tilde{H}_N(\tau;\tilde{J})] &= \mathbb{E}[\tilde{H}_{N_1}(\sigma;\tilde{J})\tilde{H}_{N_1}(\tau;\tilde{J})] + \mathbb{E}[\tilde{H}_{N_2}(\sigma;\tilde{J})\tilde{H}_{N_2}(\tau;\tilde{J})] = \\ &= N_1\langle\tilde{\mathbf{q}}^{(1)}(\sigma,\tau),\Delta^2\tilde{\mathbf{q}}^{(1)}(\sigma,\tau)\rangle + N_2\langle\tilde{\mathbf{q}}^{(2)}(\sigma,\tau),\Delta^2\tilde{\mathbf{q}}^{(2)}(\sigma,\tau)\rangle\end{aligned} \tag{4.7}$$

to be compared with (4.3). We have denoted by $\tilde{\mathbf{q}}^{(1)}$ and $\tilde{\mathbf{q}}^{(2)}$ the overlaps computed with the spins of the first and second partition respectively. By definition we can write:

$$Nq_r(\sigma,\tau) = N_1 q_r^{(1)}(\sigma,\tau) + N_2 q_r^{(2)}(\sigma,\tau) \tag{4.8}$$

Hence (4.3) becomes:

$$\begin{aligned}&\mathbb{E}[H_N(\sigma;J)H_N(\tau;J)] = \\ &\quad = N\left\langle\frac{N_1}{N}\tilde{\mathbf{q}}^{(1)}(\sigma,\tau) + \frac{N_2}{N}\tilde{\mathbf{q}}^{(2)}(\sigma,\tau),\ \Delta^2\left(\frac{N_1}{N}\tilde{\mathbf{q}}^{(1)}(\sigma,\tau) + \frac{N_2}{N}\tilde{\mathbf{q}}^{(2)}(\sigma,\tau)\right)\right\rangle\end{aligned} \tag{4.9}$$

Let us now define the interpolating hamiltonian and pressure:

$$H_N(t) = \sqrt{t}H_N^{dms} + \sqrt{1-t}\tilde{H}_N \tag{4.10}$$

$$P_N(t) = \mathbb{E}\sum_{\sigma\in\Sigma_N}e^{-\beta H_N(t)} \tag{4.11}$$

The first derivative of the pressure is:

$$
\begin{aligned}
P_N'(t) &= -\frac{\beta}{2}\mathbb{E}\omega_{N,t}\left[\frac{1}{\sqrt{t}}H_N^{dms}(\sigma;J) - \frac{1}{\sqrt{1-t}}\tilde{H}_N(\sigma;\tilde{J})\right] = \\
&= \frac{N\beta^2}{2}\mathbb{E}\Omega_{N,t}^{(2)}\Bigg[1 - \langle\tilde{\mathbf{q}}(\sigma,\tau),\Delta^2\tilde{\mathbf{q}}(\sigma,\tau)\rangle - \frac{N_1}{N} + \frac{N_1}{N}\langle\tilde{\mathbf{q}}^{(1)}(\sigma,\tau),\Delta^2\tilde{\mathbf{q}}^{(1)}(\sigma,\tau)\rangle - \\
&\qquad - \frac{N_2}{N} + \frac{N_2}{N}\langle\tilde{\mathbf{q}}^{(2)}(\sigma,\tau),\Delta^2\tilde{\mathbf{q}}^{(2)}(\sigma,\tau)\rangle\Bigg] \geq 0
\end{aligned}
\tag{4.12}
$$

thanks to the convexity of the quadratic form defined by Δ^2. Use has been made of the integration by parts in Lemma 2.1.1. $\omega_{N,t}[\cdot]$ is the Boltzmann expectation w.r.t. the interpolating hamiltonian and $\Omega_{N,t}^{(2)}(\cdot)$ denotes the expectation over the two replicas as previously defined. The equation above implies:

$$
P_N(1) \geq P_N(0) \quad \Rightarrow \quad P_{N_1}^{dms} \geq P_{N_1}^{dms} + P_{N_2}^{dms} \tag{4.13}
$$

Fekete's Lemma 1.1.1 guarantees the existence of the limit. □

4.1.2 Normalization stability

As for the previous single species models we are able to prove a normalization stability of the thermodynamic limit of the pressure. However, since this is a disordered model, we cannot compute the finite size corrections and evaluate the modifications induced on them under a change of the normalization. The following result holds also for non elliptic cases, provided that the thermodynamic limit exists.

Theorem 4.1.2 (Normalization stability for disordered multi-species models). *Consider a system with the hamiltonian* (4.2)*. If the thermodynamic limit of* (4.5) *exists, then the hamiltonian:*

$$
\tilde{H}_N^{dms}(\sigma;J) = -\frac{1}{\sqrt{N+c}}\sum_{r,s=1}^{K}\sum_{(i,j)\in\Lambda_r\times\Lambda_s} J_{ij}^{rs}\sigma_i\sigma_j \quad c\in\mathbb{R} \tag{4.14}
$$

induces the same pressure in the thermodynamic limit, namely:

$$
\lim_{N\to\infty}|p_N^{dms} - \tilde{p}_N| = 0 \tag{4.15}
$$

$$
\tilde{p}_N(\beta,h) = \frac{1}{N}\mathbb{E}\log\left[\sum_{\sigma\in\Sigma_N}\exp\left(-\beta\tilde{H}_N^{dms}(\sigma;J)\right)W_N(\sigma;\beta,\mathbf{h})\right] \tag{4.16}
$$

where W_N is defined in (4.6)*.*

Proof. Define the following interpolating hamiltonian and pressure:

$$- H_N(t;\sigma) = \frac{1}{\sqrt{N+tc}} \sum_{r,s=1}^{K} \sum_{(i,j)\in\Lambda_r\times\Lambda_s} J_{ij}^{rs}\sigma_i\sigma_j \tag{4.17}$$

$$p_N(t) = \frac{1}{N}\mathbb{E}\log\left[\sum_{\sigma\in\Sigma_N} \exp\left(-\beta H_N(t;\sigma)\right) W_N(\sigma;\beta,\mathbf{h})\right] \tag{4.18}$$

It suffices to prove that the integral of the first derivative of the interpolation pressure approaches 0 when $N\to\infty$.

$$p_N'(t) = \frac{\beta}{2N}\frac{c}{N+tc}\mathbb{E}\omega_{N,t}\left[\frac{1}{\sqrt{N+tc}} \sum_{r,s=1}^{K} \sum_{(i,j)\in\Lambda_r\times\Lambda_s} J_{ij}^{rs}\sigma_i\sigma_j\right] =$$

$$= -\frac{\beta}{2N}\frac{c}{N+tc}\mathbb{E}\omega_{N,t}\left[H_N(t;\sigma)\right] = \frac{\beta^2 c}{2N(N+tc)}\mathbb{E}\Omega_{N,t}^{(2)}\left[C(t;\sigma,\sigma) - C(t;\sigma,\tau)\right] \tag{4.19}$$

with:

$$C(t;\sigma,\tau) = \mathbb{E}[H_N(t;\sigma)H_N(t;\tau)] = \frac{N^2}{N+tc}\langle\tilde{\mathbf{q}}(\sigma,\tau),\Delta^2\tilde{\mathbf{q}}(\sigma,\tau)\rangle \tag{4.20}$$

Inserting it in the derivative we get:

$$|p_N'(t)| = \frac{N\beta^2 c}{2(N+tc)^2}|\mathbb{E}\Omega_{N,t}^{(2)}[\langle\tilde{\mathbf{1}},\Delta^2\tilde{\mathbf{1}}\rangle - \langle\tilde{\mathbf{q}}(\sigma,\tau),\Delta^2\tilde{\mathbf{q}}(\sigma,\tau)\rangle]| \le \frac{N\beta^2 cK(\alpha,\Delta)}{(N+tc)^2} \tag{4.21}$$

$$K(\alpha,\Delta) = \|\tilde{\mathbf{1}}\|^2\|\Delta^2\| \tag{4.22}$$

where $(\tilde{\mathbf{1}})_r = \alpha_r$. At this point we are finally able to prove that:

$$\left|\int_0^1 dt\, p_N'(t)\right| \le K(\alpha,\Delta)N\beta^2 c\int_0^1 dt\,\frac{1}{(N+tc)^2} = \frac{Nc\beta^2 K(\alpha,\Delta)}{N(N+c)} \longrightarrow 0 \tag{4.23}$$

when $N\to\infty$. An elementary use of the theorem of integral calculus yields:

$$p_N(1) - p_N(0) = p_N^{dms} - \tilde{p}_N = \int_0^1 dt\, p_N'(t) \longrightarrow 0 \tag{4.24}$$

when $N\to\infty$ and this concludes the proof. □

4.2 RSB ansatz for elliptic models

As in the introduction of the Parisi ansatz, we first need some definitions to generalize and adapt it to the current multi-species case. Here we follow the recursion relations presented in [19]. Consider the following non decreasing sequences:

$$0 = m_0 \le m_1 \le \cdots \le m_r \le m_{r+1} = 1 \tag{4.25}$$

$$0 = q_0^{(s)} \le q_1^{(s)} \le \cdots \le q_{r-1}^{(s)} \le q_r^{(s)} = 1 \quad s = 1, \dots, K \tag{4.26}$$

This time we denote by r the number of steps of the RSB, for K is now the number of species. We can rearrange q's in a vector as follows: $\mathbf{q}_l = (q_l^{(1)}, q_l^{(2)}, \dots, q_l^{(K)})$ and define:

$$Q_l^{(s)} = 2\Pi_s(\Delta^2 \tilde{\mathbf{q}}_l) \quad \text{with } \tilde{q}_l^{(s)} = \alpha_s q_l^{(s)} \tag{4.27}$$

Π_s is the projector on the s species. Let us now introduce the modified recursive relation:

$$Z_r^{(s)} = \cosh\left[\beta\left(h_s + \sum_{l=1}^{r} \eta_l \sqrt{Q_l^{(s)} - Q_{l-1}^{(s)}}\right)\right] \quad \eta_l \overset{\text{iid}}{\sim} \mathcal{N}(0,1) \tag{4.28}$$

$$(Z_{l-1}^{(s)})^{m_l} = \mathbb{E}_l[(Z_l^{(s)})^{m_l}] \quad \mathbb{E}_l[\cdot] = \int_{\mathbb{R}} \frac{d\eta_l}{\sqrt{2\pi}} e^{-\eta_l^2/2} \tag{4.29}$$

Definition 4.2.1 (Disordered multi-species RSB *ansatz*)**.** Given $r, m_s, q_l^{(s)}$ one can define the following ansatz functional:

$$\mathcal{P}(x;\beta,\mathbf{h}) = \log 2 + \sum_{s=1}^{K} \alpha_s \log Z_0^{(s)} - \frac{\beta^2}{2} \sum_{l=1}^{r} m_l \left(\langle \tilde{\mathbf{q}}_l, \Delta^2 \tilde{\mathbf{q}}_l \rangle - \langle \tilde{\mathbf{q}}_{l-1}, \Delta^2 \tilde{\mathbf{q}}_{l-1} \rangle\right) \tag{4.30}$$

We have already given the discrete representation of the functional, which is useful for us. If the reader is interested in a functional of a generic distribution, he will find it in [7]. As discussed in the SK model, it is sufficient to discuss the discrete case ([26]).

In this chapter we want to prove the following theorem:

Theorem 4.2.1. *The pressure in the thermodynamic limit of the disordered multi-species elliptic model is:*

$$\lim_{N\to\infty} p_N^{dms}(\beta,\mathbf{h}) = \inf \mathcal{P}(x;\beta,\mathbf{h}) \tag{4.31}$$

where the infimum is taken w.r.t. the triple $r, \{m_l\}_{1\le l\le r}, \{q_l^{(s)}\}_{1\le l\le r}^{1\le s\le K}$.

In order to find at least an upper bound we have to build a mathematical machinery similar to that of the SK model.

4.3 ROSt-cavity perspective

The aim of this section is to introduce the correct cavity functional that will allow us to state an extended variational principle, analogous to that of the SK model. This will be possible for elliptic models.

Let us follow again the guidelines by Aizenman, Sims and Starr in [3]. Imagine to have a system of N spins, α_i, $i = 1, \dots, N$ with $\alpha \in \Sigma_N$. Then add M spins σ_j, $j = 1, \dots, M$ with $\sigma \in \Sigma_M$ to it. Compare now the pressures of the system before and after the addition as follows:

$$
\begin{aligned}
G_M(\beta, \mathbf{h}) = \frac{1}{M}(P_{N+M} - P_N) = \frac{1}{M}\mathbb{E}\log\frac{Z_{N+M}}{Z_N} = \\
= \frac{1}{M}\mathbb{E}\log\frac{\sum_{\sigma,\alpha} e^{-\beta H^{dms}_{N+M}(\sigma,\alpha;J)} W_{N+M}(\sigma,\alpha;\beta,\mathbf{h})}{\sum_{\alpha} e^{-\beta H^{dms}_{N}(\alpha;J)} W_N(\alpha;\beta,\mathbf{h})}
\end{aligned} \tag{4.32}
$$

$G_M(\beta, \mathbf{h})$ is not yet the cavity functional, it is just a symbol for the moment. The N particles hamiltonian can be substituted thanks to the following equality in distribution:

$$
\begin{aligned}
-H^{dms}_N(\alpha; J) \overset{\mathrm{D}}{=} \frac{1}{\sqrt{N+M}} \sum_{r,s=1}^{K} \sum_{(i,j)\in\Lambda_r\times\Lambda_s} J^{rs}_{ij}\alpha_i\alpha_j + \\
+ \sqrt{M}\underbrace{\frac{1}{\sqrt{N(N+M)}} \sum_{r,s=1}^{K} \sum_{(i,j)\in\Lambda_r\times\Lambda_s} \tilde{J}^{rs}_{ij}\alpha_i\alpha_j}_{K_\alpha}
\end{aligned} \tag{4.33}
$$

while the $N + M$ particles hamiltonian can be split in three parts, containing: the interaction among the α's, the interaction between α's and σ's, and the interaction among σ's.

$$
\begin{aligned}
-H_{N+M}(\sigma, \alpha; J) = \frac{1}{\sqrt{M+N}} \sum_{r,s=1}^{K} \sum_{(i,j)\in\Lambda_r\times\Lambda_s} J^{rs}_{ij}\alpha_i\alpha_j + \\
+ \sum_{s=1}^{K}\sum_{j\in\Lambda_s} \underbrace{\left(\frac{\sqrt{2}}{\sqrt{M+N}} \sum_{r=1}^{K}\sum_{i\in\Lambda_r} J^{rs}_{ij}\alpha_i\right)}_{\eta^{(s)}_{j,\alpha}} \sigma_j + \frac{1}{\sqrt{M+N}} \sum_{r,s=1}^{K} \sum_{(i,j)\in\Lambda_r\times\Lambda_s} J^{rs}_{ij}\sigma_i\sigma_j
\end{aligned} \tag{4.34}
$$

When an index, say i, is referred to a spin α in a species s, it runs only over the N_s values relative to the $\alpha's$ in Λ_s. The same thing is valid for the σ's in a species

s: their index takes only M_s values. The $\sqrt{2}$ factor in the η's is due to the fact that the interaction is symmetric. We could equivalently eliminate it including a normalization factor $1/\sqrt{2(N+M)}$. Thus we have two random contributions that we put together.

If we compute the covariances of the fields K_α and $\eta^{(s)}_{j,\alpha}$ we get:

$$\mathbb{E}[K_\alpha K_{\alpha'}] = \frac{N}{N+M}\langle \tilde{\mathbf{p}}(\alpha,\alpha'), \Delta^2 \tilde{\mathbf{p}}(\alpha,\alpha')\rangle \tag{4.35}$$

$$\mathbb{E}[\eta^{(s)}_{j,\alpha}\eta^{(s')}_{j',\alpha'}] = \frac{2N}{N+M}\delta_{jj'}\delta_{ss'}\Pi_s(\Delta^2\tilde{\mathbf{p}}(\alpha,\alpha')) \tag{4.36}$$

Π_s is the orthogonal projector on the species s and $\tilde{p}_r(\alpha,\alpha') = \alpha_r p_r(\alpha,\alpha')$, where $p_r(\alpha,\alpha')$ is the overlap of the spins α in the species r . Introducing the following notation:

$$\xi_\alpha = \exp\left(\frac{\beta}{\sqrt{N+M}}\sum_{r,s=1}^{K}\sum_{(i,j)\in\Lambda_r\times\Lambda_s} J^{rs}_{ij}\alpha_i\alpha_j + \beta\sum_{r=1}^{K} h_r \sum_{i\in\Lambda_r}\alpha_i\right) \tag{4.37}$$

we can write the quantity G_M as follows:

$$G_M(\beta,\mathbf{h}) = \frac{1}{M}\mathbb{E}\log\left[\frac{\sum_{\alpha,\sigma}\xi_\alpha \exp\left(\beta\sum_{s=1}^{K}\sum_{j\in\Lambda_s}(\eta^{(s)}_{j,\alpha}+h_s)\sigma_j\right)}{\sum_\alpha \xi_\alpha \exp\left(\beta\sqrt{M}K_\alpha\right)}\right] \tag{4.38}$$

We are finally led to the following definition of the cavity functional for a multi-species model.

Definition 4.3.1 (Cavity functional, disordered multi-species models)**.** Let $r = (\xi,\mathbf{p})$ be a ROSt. Consider the gaussian families K_α and $\eta^{(s)}_{j,\alpha}$ with:

$$\mathbb{E}[K_\alpha K_{\alpha'}] = \langle \tilde{\mathbf{p}}(\alpha,\alpha'), \Delta^2 \tilde{\mathbf{p}}(\alpha,\alpha')\rangle \tag{4.39}$$

$$\mathbb{E}[\eta^{(s)}_{j,\alpha}\eta^{(s')}_{j',\alpha'}] = 2\delta_{jj'}\delta_{ss'}\Pi_s(\Delta^2\tilde{\mathbf{p}}(\alpha,\alpha')) \tag{4.40}$$

The cavity functional is then defined as:

$$G_{r.M}(\beta,\mathbf{h}) = \frac{1}{M}\mathbb{E}\log\left[\frac{\sum_{\alpha,\sigma}\xi_\alpha \exp\left(\beta\sum_{s=1}^{K}\sum_{j\in\Lambda_s}(\eta^{(s)}_{j,\alpha}+h_s)\sigma_j\right)}{\sum_\alpha \xi_\alpha \exp\left(\beta\sqrt{M}K_\alpha\right)}\right] \tag{4.41}$$

Remark 4.3.1. We are basically reproducing the ideas shown in [19] by Panchenko in the language of the cavity functional. In fact, the covariances (2.58) and (2.57) remind us of the covariances (18) in the mentioned paper.

4.3.1 Extended variational principle

As for the SK model we have to prove that the pressure of the model is bounded both from below and above by $\inf_r G_{r,M}(\beta, \mathbf{h})$.

Proposition 4.3.1. *$\forall\, M \in \mathbb{N}$ and for any ROSt r the following inequality holds for an elliptic model:*

$$p_M^{dms}(\beta, \mathbf{h}) \le G_{r,M}(\beta, \mathbf{h}) \tag{4.42}$$

Optimizing w.r.t. the possibile ROSts we get:

$$p_M^{dms}(\beta, \mathbf{h}) \le \inf_r G_{r,M}(\beta, \mathbf{h}) \tag{4.43}$$

Proof. Again the validity of the equality is guaranteed through convexity arguments. Considering p_M^{dms} as the one defined in (4.5), with N replaced by M, the statement can be recast in the following form:

$$\mathbb{E}\log\left[\sum_{\sigma,\alpha}\xi_\alpha W_M(\sigma;\beta,\mathbf{h})\left(\beta\left(\frac{1}{\sqrt{M}}\sum_{r,s=1}^{K}\sum_{(i,j)\in\Lambda_r\times\Lambda_s}J_{ij}^{rs}\sigma_i\sigma_j+\sqrt{M}K_\alpha\right)\right)\right]\le$$

$$\le \mathbb{E}\log\left[\sum_{\alpha,\sigma}\xi_\alpha W_M(\sigma;\beta,\mathbf{h})\exp\left(\beta\sum_{s=1}^{K}\sum_{j\in\Lambda_s}\eta_{j,\alpha}^{(s)}\sigma_j\right)\right] \tag{4.44}$$

Consider now the following two families, together with their covariances:

$$A_{\sigma,\alpha} = \frac{1}{\sqrt{M}}\sum_{r,s=1}^{K}\sum_{(i,j)\in\Lambda_r\times\Lambda_s}J_{ij}^{rs}\sigma_i\sigma_j+\sqrt{M}K_\alpha \tag{4.45}$$

$$B_{\sigma,\alpha} = \sum_{s=1}^{K}\sum_{j\in\Lambda_s}(\eta_{j,\alpha}^{(s)}+h_s)\sigma_j \tag{4.46}$$

$$\mathbb{E}[A_{\sigma,\alpha}A_{\sigma',\alpha'}] = M\langle\tilde{\mathbf{q}}(\sigma,\sigma'),\Delta^2\tilde{\mathbf{q}}(\sigma,\sigma')\rangle + M\langle\tilde{\mathbf{p}}(\alpha,\alpha'),\Delta^2\tilde{\mathbf{p}}(\alpha,\alpha')\rangle \tag{4.47}$$

$$\mathbb{E}[B_{\sigma,\alpha}B_{\sigma',\alpha'}] = 2M\langle\tilde{\mathbf{q}}(\sigma,\sigma'),\Delta^2\tilde{\mathbf{p}}(\alpha,\alpha')\rangle \tag{4.48}$$

To compute the covariances we have used the definitions (4.39) and (4.40). Notice that:

$$\mathbb{E}[A_{\sigma,\alpha}A_{\sigma,\alpha}] = \mathbb{E}[B_{\sigma,\alpha}B_{\sigma,\alpha}] \tag{4.49}$$

$$\mathbb{E}[A_{\sigma,\alpha}A_{\sigma',\alpha'}] - \mathbb{E}[B_{\sigma,\alpha}B_{\sigma',\alpha'}] = M\langle\tilde{\mathbf{q}}(\sigma,\sigma')-\tilde{\mathbf{p}}(\alpha,\alpha'),\Delta^2(\tilde{\mathbf{q}}(\sigma,\sigma')-\tilde{\mathbf{p}}(\alpha,\alpha'))\rangle \ge 0 \tag{4.50}$$

because Δ^2 is positive definite. Finally, thanks again to the comparison between gaussian families (Theorem 3.46 in [10]) the proposition is proved. □

The following result, provides a lower bound for the extended variational principle. Notice that it holds even for hyperbolic models.

Proposition 4.3.2. *There exists a ROSt r such that:*

$$\liminf_{N\to\infty} \frac{P^{dms}_{N+M} - P^{dms}_{N}}{M} = G_{r,M}(\beta, \mathbf{h}) \tag{4.51}$$

where P^{dms}_N is the extensive pressure of a multi-species model.

Proof. We have already proved this result implictly in the construction of the cavity functional. In fact the fields (4.35) and (4.36), if we perform the $\liminf_{N\to\infty}$, satisfy exactly the relations (4.39) and (4.40) respectively. We would have an extra term in the exponent containing interactions between the σ's only. In the $\liminf_{N\to\infty}$ that very term is irrelevant. □

Theorem 4.3.3 (Extended variational principle, multi-species models)**.** *The pressure of a disordered multi-species elliptic model is:*

$$\lim_{N\to\infty} p^{dms}_N = \lim_{M\to\infty} \inf_r G_{r,m}(\beta, \mathbf{h}) \tag{4.52}$$

Proof. Thanks to Lemma 2.3.6, and to the previous proposition, there exists a ROSt r such that:

$$\liminf_{N\to\infty} \frac{P^{dms}_N}{N} \geq \liminf_{N\to\infty} \frac{P^{dms}_{N+M} - P^{dms}_{N}}{M} = G_{r,M}(\beta, \mathbf{h}) \geq \inf_r G_{r,M}(\beta, \mathbf{h}) \tag{4.53}$$

Sending $M \to \infty$ we get the correct lower bound.

The upper bound is instead guaranteed by the fact that:

$$p^{dms}_M(\beta, \mathbf{h}) \leq \inf_r G_{r,M}(\beta, \mathbf{h}) \tag{4.54}$$

Sending $M \to \infty$ we get the upper bound and this concludes the proof. □

4.4 Upper bound via cavity functional

In order to find an upper bound, which will be uniform in the number of spins, and thus equal to our ansatz, we have to exhibit a ROSt $\bar{r}$ such that:

$$\mathcal{P}(x; \beta, \mathbf{h}) = G_{\bar{r},M}(\beta, \mathbf{h}) \tag{4.55}$$

The comparison between the cavity functional and the pressure of the system then does the rest. The solution is again to build a ROSt based on Ruelle probability cascades, as stated and proved in the following theorem.

Theorem 4.4.1. *Consider a ROSt* $\bar{r} = (\bar{\xi}, \bar{\mathbf{p}})$*, where* $\bar{\xi}_\alpha$ *are RPCs random weights as in* (2.47) *and* $p^{(s)}_{\alpha,\alpha'}$ *has an ultrametric structure for* $s = 1, \ldots, K$*.*

$$p^{(s)}_{\alpha,\alpha'} = \begin{cases} q^{(s)}_0 & \textit{if } \alpha_1 \neq \alpha'_1 \\ q^{(s)}_1 & \textit{if } \alpha_1 = \alpha'_1 \textit{ and } \alpha_2 \neq \alpha'_2 \\ \vdots & \\ q^{(s)}_r = 1 & \textit{if } \alpha_i = \alpha'_i \; \forall\, i = 1, \ldots, r \end{cases} \tag{4.56}$$

Then we can represent the ansatz (4.30) *as follows:*

$$\mathcal{P}(x; \beta, \mathbf{h}) = G_{\bar{r},M}(\beta, \mathbf{h}) \tag{4.57}$$

Proof. The key is again the invariance properties of PPP (2.46), inherited by the random weights $\bar{\xi}_\alpha$.

First, we have to find two explicit forms of the fields K_α and $\eta^{(s)}_{j,\alpha}$ that satisfy (4.39) and (4.40) respectively, where the overlap matrix is the ultrametric ansatz $\bar{\mathbf{p}}$. This time the choice is a little more tricky. The following field:

$$\begin{aligned}\bar{K}_\alpha = \sqrt{\langle \tilde{\mathbf{q}}_1, \Delta^2 \tilde{\mathbf{q}}_1 \rangle - \langle \tilde{\mathbf{q}}_0, \Delta^2 \tilde{\mathbf{q}}_0 \rangle}\, \tilde{J}_{\alpha_1} + \sqrt{\langle \tilde{\mathbf{q}}_2, \Delta^2 \tilde{\mathbf{q}}_2 \rangle - \langle \tilde{\mathbf{q}}_1, \Delta^2 \tilde{\mathbf{q}}_1 \rangle}\, \tilde{J}_{\alpha_1\alpha_2} + \ldots \\ \cdots + \sqrt{\langle \tilde{\mathbf{q}}_r, \Delta^2 \tilde{\mathbf{q}}_r \rangle - \langle \tilde{\mathbf{q}}_{r-1}, \Delta^2 \tilde{\mathbf{q}}_{r-1} \rangle}\, \tilde{J}_{\alpha_1\alpha_2\ldots\alpha_r}\end{aligned} \tag{4.58}$$

where the $\tilde{J}$'s are all independent gaussian centered.

$$\mathbb{E}[K_\alpha K_{\alpha'}] = \langle \tilde{\mathbf{q}}_l, \Delta^2 \tilde{\mathbf{q}}_l \rangle \quad \text{with } l = min\{0 \leq l \leq r | \alpha_1 = \alpha'_1, \ldots, \alpha_l = \alpha'_l, \alpha_{l+1} \neq \alpha'_{l+1}\} \tag{4.59}$$

which is nothing but $\langle \tilde{\mathbf{p}}_{\alpha,\alpha'}, \Delta^2 \tilde{\mathbf{p}}_{\alpha,\alpha'} \rangle$. The condition (4.39) is fulfilled. A similar choice can be made for the remaining fields:

$$\eta^{(s)}_{j,\alpha} = \sqrt{Q^{(s)}_1 - Q^{(s)}_0}\, J^{(s)}_{j,\alpha_1} + \sqrt{Q^{(s)}_2 - Q^{(s)}_1}\, J^{(s)}_{j,\alpha_1\alpha_2} + \cdots + \sqrt{Q^{(s)}_r - Q^{(s)}_{r-1}}\, J^{(s)}_{j,\alpha_1\ldots\alpha_r} \tag{4.60}$$

Again the J's are all independent gaussian centered. In such a way a simple computation yields:

$$\mathbb{E}[\eta^{(s)}_{j,\alpha} \eta^{(s')}_{j',\alpha'}] = \delta_{jj'} \delta_{ss'} Q^{(s)}_l = 2 \delta_{jj'} \delta_{ss'} \Pi_s \left(\Delta^2 \tilde{\mathbf{q}}_l \right) \tag{4.61}$$

with $l = min\{0 \leq l \leq r | \alpha_1 = \alpha'_1, \ldots, \alpha_l = \alpha'_l, \alpha_{l+1} \neq \alpha'_{l+1}\}$. Hence (4.40) is fulfilled.

Now, let us analyze the contribution of the denominator and numerator in $G_{\bar{r},M}$ separately. For the numerator we have:

$$
\begin{aligned}
&\frac{1}{M}\mathbb{E}\log\left[\sum_{\sigma,\alpha}\bar{\xi}_\alpha\exp\left(\beta\sum_{s=1}^{K}\sum_{j\in\Lambda_s}(\eta_{j,\alpha}^{(s)}+h_s)\sigma_j\right)\right]=\\
&\qquad=\frac{1}{M}\mathbb{E}\log\left[\sum_{\alpha}\bar{\xi}_\alpha\prod_{s=1}^{K}\prod_{j\in\Lambda_s}\exp\log 2\cosh\beta\left(\eta_{j,\alpha}^{(s)}+h_s\right)\right]=\\
&=\frac{1}{M}\mathbb{E}\log\left[\sum_{\alpha}\bar{\xi}_\alpha\prod_{s=1}^{K}\prod_{j\in\Lambda_s}\mathbb{E}_{j,r}^{1/m_r}\left(\exp m_r\log 2\cosh\beta\left(\sqrt{Q_r^{(s)}-Q_{r-1}^{(s)}}\eta_{j,r}+\dots\right.\right.\right.\\
&\qquad\qquad\left.\left.\left.\dots+\sqrt{Q_1^{(s)}-Q_0^{(s)}}J_{j,\alpha_1}^{(s)}+h_s\right)\right)\right]=\dots\\
&\qquad\qquad\qquad\dots=\log 2+\sum_{s=1}^{K}\alpha_s\log Z_0^{(s)}+\mathbb{E}\log\sum_\alpha\xi_\alpha
\end{aligned}
\tag{4.62}
$$

Finally for the denominator:

$$
\begin{aligned}
&\frac{1}{M}\mathbb{E}\log\sum_\alpha\bar{\xi}_\alpha\exp\left(\beta\sqrt{M}K_\alpha\right)=\\
&=\frac{1}{M}\mathbb{E}\log\sum_\alpha\bar{\xi}_\alpha\mathbb{E}_r^{1/m_r}\left[\exp\left(m_r(\beta\sqrt{M}(\sqrt{\langle\tilde{\mathbf{q}}_1,\Delta^2\tilde{\mathbf{q}}_1\rangle-\langle\tilde{\mathbf{q}}_0,\Delta^2\tilde{\mathbf{q}}_0\rangle}\tilde{J}_{\alpha_1}+\dots\right.\right.\\
&\qquad\qquad\left.\left.\dots+\sqrt{\langle\tilde{\mathbf{q}}_r,\Delta^2\tilde{\mathbf{q}}_r\rangle-\langle\tilde{\mathbf{q}}_{r-1},\Delta^2\tilde{\mathbf{q}}_{r-1}\rangle}\eta_r))\right)\right]=\\
&\quad=\frac{1}{M}\mathbb{E}\log\sum_\alpha\bar{\xi}_\alpha\exp\left(\beta\sqrt{M}(\sqrt{\langle\tilde{\mathbf{q}}_1,\Delta^2\tilde{\mathbf{q}}_1\rangle-\langle\tilde{\mathbf{q}}_0,\Delta^2\tilde{\mathbf{q}}_0\rangle}\tilde{J}_{\alpha_1}+\dots\right.\\
&\qquad\qquad\left.\dots+\sqrt{\langle\tilde{\mathbf{q}}_{r-1},\Delta^2\tilde{\mathbf{q}}_{r-1}\rangle-\langle\tilde{\mathbf{q}}_{r-2},\Delta^2\tilde{\mathbf{q}}_{r-2}\rangle}\tilde{J}_{\alpha_1\dots\alpha_{r-1}})\right)+\\
&\qquad\qquad+\frac{\beta^2}{2}m_r(\langle\tilde{\mathbf{q}}_r,\Delta^2\tilde{\mathbf{q}}_r\rangle-\langle\tilde{\mathbf{q}}_{r-1},\Delta^2\tilde{\mathbf{q}}_{r-1}\rangle)=\dots\\
&\qquad\qquad\dots=\mathbb{E}\log\sum_\alpha\bar{\xi}_\alpha+\frac{\beta^2}{2}\sum_{l=1}^{r}m_l\left(\langle\tilde{\mathbf{q}}_l,\Delta^2\tilde{\mathbf{q}}_l\rangle-\langle\tilde{\mathbf{q}}_{l-1},\Delta^2\tilde{\mathbf{q}}_{l-1}\rangle\right)
\end{aligned}
\tag{4.63}
$$

We have used the fact that:

$$
\mathbb{E}[e^{tX}]=e^{\frac{t^2\sigma^2}{2}}\quad\text{with } X\sim\mathcal{N}(0,\sigma^2).
\tag{4.64}
$$

In both cases the dots indicate the repeated use of the invariance property (2.46). Subtracting the second contribution from the first one we get:

$$G_{\bar{r},M}(\beta,\mathbf{h}) = \log 2 + \sum_{s=1}^{K} \alpha_s \log Z_0^{(s)} - \frac{\beta^2}{2} \sum_{l=1}^{r} m_l \left(\langle \tilde{\mathbf{q}}_l, \Delta^2 \tilde{\mathbf{q}}_l \rangle - \langle \tilde{\mathbf{q}}_{l-1}, \Delta^2 \tilde{\mathbf{q}}_{l-1} \rangle \right) \tag{4.65}$$

□

Corollary 4.4.2 (Upper bound for the pressure, disordered multi-species elliptic model)**.**

$$p_N^{dms}(\beta,\mathbf{h}) \leq \mathcal{P}(x;\beta,\mathbf{h}) \tag{4.66}$$

uniformly in N*. Optimizing with respect to the triple* $r, \{m_l\}_{1\leq l\leq r}, \{q_l^{(s)}\}_{1\leq l\leq r}^{1\leq s\leq K}$ *we get:*

$$p_N^{dms}(\beta,\mathbf{h}) \leq \inf \mathcal{P}(x;\beta,\mathbf{h}) \tag{4.67}$$

Proof. The proof follows immediately from the fact that $p_M^{dms} \leq G_{r,M}$ for any ROSt r and from the previous representation theorem. □

4.5 Hints for the lower bound

As we have already done for SK, here we list a series of results used to obtain a lower bound ([19]). The procedure is basically the same, except for slight changes due to the presence of more than one species.

To begin with, we need to prove that some Ghirlanda-Guerra identities hold for the limiting array of overlaps $q_{s,ll'}$, to which the sequences $q_{s,N}(\sigma^l,\sigma^{l'})$ of the finite size overlaps converge weakly. To this purpose, this time we have to introduce a further parameter $w \in W$, with W a dense countable subset of $[0,1]^K$, and a perturbing hamiltonian:

$$g(\sigma) = \sum_{w\in W} \sum_{s=1}^{K} \frac{x_{w,p}}{2^{j(w)+p}} g_{N,w,p}(\sigma) \tag{4.68}$$

$$\mathbb{E}\left[g_{N,w,p}(\sigma^l) g_{N,w,p}(\sigma^{l'}) \right] = \underbrace{\left(\sum_{s=1}^{K} \alpha_s w_s q_s(\sigma^l,\sigma^{l'}) \right)^p}_{q_w^p(\sigma^l,\sigma^{l'})} \tag{4.69}$$

where $j(w)$ is a bijection from the countable set W to $\mathbb{N}$, and $x_{w,p}$ are uniformly distributed r.v. in $[1,2]$. Similarly to SK case, this perturbation is added to the hamiltonian (4.2) with a sequence that reaches 0 when $N \to \infty$ appropriately: $H_N^{dms}(\sigma) + s_N g(\sigma)$. If we denote by $\langle \cdot \rangle$ the Gibbs random measure induced by the perturbed hamiltonian, the following theorem holds [19]:

Theorem 4.5.1. *For any $n \geq 2$ (number of replicas), $p \geq 1$ and for any function f of the overlaps:*

$$\lim_{N\to\infty} \mathbb{E}_x \left[\mathbb{E}\langle f q^p_{w,1,n+1} \rangle - \frac{1}{n}\mathbb{E}\langle f \rangle \mathbb{E}\langle q^p_{w,12}\rangle - \frac{1}{n}\sum_{l=2}^{n} \mathbb{E}\langle f q^p_{w,1l}\rangle \right] = 0 \tag{4.70}$$

Hence, there exists again a deterministic sequence $x_{N,w,p}$ that allows us to eliminate the average. Recall that, for the limiting overlap matrix $q_{s,ll'}$ there's an appropriate measure that satisfies identities (2.4.4) with generic bounded functions of the overlaps, say $\phi(\mathbf{q}_{ll'})$.

Notice that the total overlap of the system may be written in this way:

$$q_N(\sigma^l, \sigma^{l'}) = \frac{1}{N}\sum_{i=1}^{N} \sigma_i^l \sigma_i^{l'} = \sum_{s=1}^{K} \alpha_s q_s(\sigma^l, \sigma^{l'}) \tag{4.71}$$

therefore the respective limiting overlaps will satisfy:

$$q_{ll'} = \sum_{s=1}^{K} \alpha_s q_{s,ll'} \tag{4.72}$$

Before employing the Aizenmann-Sims-Starr scheme to get the lower bound we need a synchronization theorem ([19]):

Theorem 4.5.2 (Synchronization of species)**.** *For limiting overlaps $q_{s,ll'}$ satisfying the identities* (2.4.4) *, there exist deterministic and non decreasing functions $1/\alpha_s$-Lipschitz functions $L_s : [0,1] \longrightarrow [0,1]$ such that $L_s(q_{ll'}) = q_{s,ll'}$, a.s. $\forall\, s = 1, \dots, K,\ l, l' \geq 1$.*

This theorem will allow us to use only one RPC sequence to approximate the distribution of the whole system overlap. Then, the single species overlap can be derived through the corresponding Lipschitz functions.

In order to find the lower bound we exploit the following inequality, that follows again from Lemma 2.3.6:

$$\liminf_{N\to\infty} p_N^{dms} \geq \liminf_{N\to\infty} \frac{1}{M}\left[\log Z^{dms}_{N+M} - Z^{dms}_N\right] \tag{4.73}$$

We have already studied the difference in the square bracket, used to build the cavity functional for an extended variational principle. This construction may be done also with the previously introduced perturbing hamiltonian, and the result is unchanged up to irrelevant terms. In addition to that Theorem 4.5.1 applies also to the new perturbed measure $\langle \cdot \rangle'$, where the normalization constants in the hamiltonians have been changed according to the substitution: $N \to N + M$. Hence, except for the perturbed measure, the procedure yields the usual result:

$$\liminf_{N\to\infty} A_N(x) = \liminf_{N\to\infty} \frac{1}{M} \left[\mathbb{E} \log \langle \prod_{s=1}^{K} \prod_{j\in\Lambda_s} 2\cosh \eta_{j,\sigma}^{(s)} \rangle' - \mathbb{E} \log \langle \exp K(\sigma) \rangle' \right] \quad (4.74)$$

where the fields K and $\eta^{(s)}$ are basically (up to temperature factors) the ones in (4.35) and (4.36) respectively. We are neglecting the external fields for the moment, since their introduction does not change this arguments significantly.

The Ghirlanda Guerra identities are applicable also to the total asymptotic overlap $q_{ll'}$, so, by theorem (2.4.6) there's a matrix $Q_{ll'} = \sigma^l \cdot \sigma^{l'}$, with σ^l sampled according to an RPC, that approximates $q_{ll'}$ in distribution. Let the RPC parameters be:

$$0 = q_0 \le q_1 \le \cdots \le q_{r-1} \le q_r = 1 \quad (4.75)$$
$$0 = m_0 \le m_1 \le \cdots \le m_r \le m_{r+1} = 1 \quad (4.76)$$

Now, if we set $q_{s,l} = L_s(q_l)$ and $Q_{s,ll'} = L_s(Q_{ll'})$, we will have that $Q_{s,ll'}$ approximate $q_{s,ll'}$ in distribution by construction. Recall that, under the RPC measure, the overlap matrix has a distribution similar to that in (2.82).

Once we have proved that the lower bound can be written with Ruelle probability cascades, thanks to the representation of the RSB ansatz (4.30) in terms of them, the identification is possible and the proof is completed.

Chapter 5

Annealed regions and replica symmetry

Up to now we have dealt only with quenched pressures, *i.e.* with the gaussian average taken after the log is performed on the partition function. We now investigate what happens when the average is taken directly on the partition function. It turns out that in this case the pressure can be explicitly computed thanks to the gaussian distribution of the interactions. We wonder if there is a region, in the phase space (β, h), where the two pressures coincide in the thermodynamic limit. This region exists and we will call it *annealed region* or *regime*, not to be confused with the replica symmetric region, that we will discuss later, though they coincide (if the AT line is correct, see below) for the SK model with vanishing external field. This region is characterized by a high temperature. The annealed regime plays a central role in the learning of deep networks. In particular, for these purposes, this region has to be the smallest possible, because here, intuitively, the thermal noise becomes too intense. From these considerations we will get a geometric constraint on the ratios between the layers sizes of a shallow restricted Boltzmann machine (up to 4 layers). These relations are particularly simple when the variances of the interactions between the layers are equal.

Let us be more precise with some preliminary definitions.

Definition 5.0.1 (Annealed pressure)**.** Given an hamiltonian $H_N(\sigma, J)$, that contains the disorders J, the annealed pressure per particle of the corresponding model is:

$$p_N^{an}(\beta, h) = \frac{1}{N} \log \mathbb{E} \sum_{\sigma \in \Sigma_N} e^{-\beta H_N(\sigma, J)} \tag{5.1}$$

Definition 5.0.2 (Annealed region)**.** Let $H_N(\sigma, J)$ be as before. The annealed is a subset of the phase space (β, h) in which the quenched and annealed pressure of the model coincide:

$$\lim_{N\to\infty} p_N^{an}(\beta,h) = \lim_{N\to\infty} p_N(\beta,h) = p(\beta,h) \tag{5.2}$$

$$p_N(\beta,h) = \frac{1}{N}\mathbb{E}\log\sum_{\sigma\in\Sigma_N} e^{-\beta H_N(\sigma,J)} \quad \text{(quenched pressure)} \tag{5.3}$$

5.1 Annealing in the SK model

Let us begin with the simplest, mono-species case: the SK model. The we will generalize it to multi-species models. This was first discussed in [1]. In the case of vanishing external field $h = 0$ we are able to prove the following theorem. Here, we follow a strategy based on the second moment method, also used in [7]. We will need the following lemma.

Lemma 5.1.1. *Let Z be a random variable, in the probability measure $\mathbb{P}$. Then:*

$$\mathbb{P}[Z \geq \mathbb{E}[Z]/2] \geq \frac{\mathbb{E}^2[Z]}{4\mathbb{E}[Z^2]} \tag{5.4}$$

Proof. Using Cauchy-Schwartz's inequality:

$$\mathbb{E}[Z] = \mathbb{E}[Z|Z \geq \mathbb{E}[Z]/2] + \mathbb{E}[Z|Z < \mathbb{E}[Z]/2] \leq \sqrt{\mathbb{P}[Z \geq \mathbb{E}[Z]/2]\mathbb{E}[Z^2]} + \frac{\mathbb{E}[Z]}{2} \tag{5.5}$$

Solving the inequality for $\mathbb{P}$ we get the result. $\square$

Theorem 5.1.2. *The quenched and annealed pressures of the SK model coincide in the thermodynamic limit for $\beta^2 \leq 1/2$ and $h = 0$. More precisely:*

$$p^{SK}(\beta,0) = \lim_{N\to\infty}\frac{1}{N}\mathbb{E}\log\sum_{\sigma\in\Sigma_N} e^{-\beta H_N^{SK}(\sigma,J)} = \\ = \lim_{N\to\infty}\frac{1}{N}\log\mathbb{E}\sum_{\sigma\in\Sigma_N} e^{-\beta H_N^{SK}(\sigma,J)} = p^{an}(\beta,0) \tag{5.6}$$

Proof. Thanks to Jensen's inequality and to the concavity of the logarithm we can say that:

$$p_N^{SK} \leq \frac{1}{N}\log\sum_{\sigma\in\Sigma_N}\mathbb{E}e^{\frac{\beta}{\sqrt{N}}\sum_{i,j=1}^N J_{ij}\sigma_i\sigma_j} = \frac{1}{N}\log\sum_{\sigma\in\Sigma_N} e^{\frac{N\beta^2}{2}} = \log 2 + \frac{\beta^2}{2} \tag{5.7}$$

Now we proceed with a comparison between $\mathbb{E}[Z_N^2]$ e $\mathbb{E}^2[Z_N]$, checking that:

$$\frac{\mathbb{E}^2[Z_N]}{\mathbb{E}[Z_N^2]} \geq C > 0 \tag{5.8}$$

uniformly in N. Z_N is, as usual, the partition function. Let us start by computing:

$$\mathbb{E}[Z_N^2] = \sum_{\sigma,\tau\in\Sigma_N} \mathbb{E}e^{-\beta[H_\sigma^{SK}+H_\tau^{SK}]} = \sum_{\sigma,\tau\in\Sigma_N} e^{\frac{\beta^2 N}{2}[2+2q_N(\sigma,\tau)]} = e^{\frac{\beta^2 N}{2}+N\log 2}\sum_\sigma e^{N\beta^2 m_N(\sigma)} \tag{5.9}$$

In the last step we have used the gauge freedom $\sigma_i \longrightarrow \sigma_i\tau_i$, which is allowed only in the case $h = 0$. Using Hubbard-Stratonovič transform to linearize the quadratic term we get:

$$\begin{aligned}
\mathbb{E}[Z_N^2] &= e^{\frac{\beta^2 N}{2}+N\log 2}\sum_\sigma \int_{\mathbb{R}} \frac{dx}{\sqrt{2\pi\beta^2 2N}} e^{-\frac{x^2}{2\beta^2 2N}+m_N(\sigma)x} = \\
&= \frac{e^{\frac{\beta^2 N}{2}+N\log 2}}{\beta}\sum_\sigma \int_{\mathbb{R}} \frac{dx}{\sqrt{2\pi}} e^{-\frac{x^2}{2\beta^2}+\sqrt{2N}m_N(\sigma)x} = \frac{e^{\frac{\beta^2 N}{2}+2N\log 2}}{\beta}\times \\
&\times \int_{\mathbb{R}} \frac{dx}{\sqrt{2\pi}} e^{-\frac{x^2}{2\beta^2}+N\log\cosh\frac{\sqrt{2N}}{N}x} \leq \frac{e^{\frac{\beta^2 N}{2}+2N\log 2}}{\beta}\int_{\mathbb{R}} \frac{dx}{\sqrt{2\pi}} e^{-x^2\left(\frac{1}{2\beta^2}-1\right)}
\end{aligned} \tag{5.10}$$

The last inequality is valid thanks to the fact that $\log\cosh x \leq x^2/2$. The integral is convergent only if $\beta^2 < 1/2$. This implies that the constant C above is non vanishing in this very case. Lemma 5.1.1 implies that:

$$\mathbb{P}\left(\frac{1}{N}\log Z_N \geq \frac{1}{N}\log\mathbb{E}[Z_N] - \frac{\log 2}{N}\right) \geq C \tag{5.11}$$

Finally, thanks to the self-averaging property of the pressure, the following inequality must hold:

$$\lim_{N\to\infty}\frac{1}{N}\log Z_N = \lim_{N\to\infty}\frac{1}{N}\mathbb{E}\log Z_N \geq \lim_{N\to\infty}\frac{1}{N}\log\mathbb{E}[Z_N] = \log 2 + \frac{\beta^2}{2} \tag{5.12}$$

The claim for $\beta^2 = 1/2$ also holds, thanks to continuity, because p is convex in β and bounded by the annealed solution. □

Remark 5.1.1. This statement identifies a sufficient condition, hence it does not define the annealed region precisely.

With the following argument we aim to prove that if $\beta^2 > 1/2$ then the quenched and annealed pressure differ. Consider the consistency equation (2.34). For $\beta^2 > 1/2$, it gives also a positive solution, in addition to $\bar{q} = 0$, that this time is not stable. This can be easily shown by checking the the r.h.s. of the consistency equation (2.34) is no longer a contractive map in $\bar{q} = 0$. Hence the other solution now realizes a minimum of our variational replica symmetric pressure. More rigorously:

Proposition 5.1.3. *Consider the replica symmetric pressure p_{RS} evaluated in its extremal point, called $\bar{q}$, that satisfies* (2.34). *Then:*

$$p_{RS}(\beta, 0; \bar{q}) \begin{cases} < \frac{\beta^2}{2} + \log 2 & \text{if } \beta^2 > \frac{1}{2} \\ = \frac{\beta^2}{2} + \log 2 & \text{if } \beta^2 \leq \frac{1}{2} \end{cases} \tag{5.13}$$

Proof. To prove this fact, we compute the first derivative with respect to β of $p_{RS}(\beta, 0)$, keeping in mind that $\bar{q}$ is also a function of β.

$$\begin{aligned} \frac{dp_{RS}(\beta,0)}{d\beta} &= \beta(1-\bar{q})^2 + \beta^2(\bar{q}-1)\frac{d\bar{q}}{d\beta} + \mathbb{E}_z\left[\tanh\left(\beta\sqrt{2\bar{q}}z\right) z \left(\sqrt{2\bar{q}} + \frac{\beta}{\sqrt{2\bar{q}}}\frac{d\bar{q}}{d\beta}\right)\right] = \\ &= \beta(1-\bar{q})^2 + \beta^2(\bar{q}-1)\frac{d\bar{q}}{d\beta} + (1-\bar{q})\beta\left(2\bar{q} + \beta\frac{d\bar{q}}{d\beta}\right) = \beta(1-\bar{q}^2) \end{aligned} \tag{5.14}$$

If we rewrite the pressure in integral form we get:

$$p_{RS}(\beta, 0) = \log 2 + \int_0^\beta d\beta' \beta'(1 - \bar{q}^2(\beta')) \tag{5.15}$$

We immediately get the annealed pressure if $\beta^2 \leq 1/2$, beacuse $\bar{q} = 0$. Otherwise, splitting the integral:

$$\begin{aligned} p_{RS}(\beta, 0) = \log 2 + \frac{1}{4} + \int_{1/\sqrt{2}}^\beta d\beta' \beta'(1 - \bar{q}^2(\beta')) < \\ < \log 2 + \frac{1}{4} + \int_{1/\sqrt{2}}^\beta d\beta' \beta' = \log 2 + \frac{\beta^2}{2} \end{aligned} \tag{5.16}$$

The strict inequality follows from the fact that for $\beta^2 > 1/2$ $\bar{q} > 0$, except for the inferior extremum in the integral that can be neglected. $\square$

Putting this together with Theorem 5.1.2 and Guerra's replica symmetric bound, we get the following corollary, that fully characterizes the annealed region for SK with vanishing external field.

Corollary 5.1.4. *For $h = 0$, $\beta^2 \leq \frac{1}{2}$ are all and only the possible values of inverse absolute temperature for which the annealed pressure equals the quenched pressure in the thermodynamic limit.*

Proof. It simply follows from the fact that:

$$p_N^{SK}(\beta, 0) \leq p_{RS}(\beta, 0; \bar{q}) < \log 2 + \frac{\beta^2}{2} = p^A(\beta, 0) \quad \text{when } \beta^2 > \frac{1}{2} \tag{5.17}$$

whereas:

$$p_{RS}(\beta, 0; \bar{q} = 0) = \log 2 + \frac{\beta^2}{2} = p^A(\beta, 0) \quad \text{when } \beta^2 \leq \frac{1}{2} \tag{5.18}$$

□

5.1.1 The Almeida-Thouless line (AT)

For the SK model, the annealed region is strongly related to the replica symmetric region, where $p^{SK} = p_{RS}$. It turns out that it is very difficult to characterize the latter. However, it is clear that there is a line in the phase space, the Almeida-Thouless line, beyond which the replica symmetric pressure cannot be the true pressure of the model ([27]).

Theorem 5.1.5 (Almeida-Thouless line)**.** *If:*

$$\beta^2 \mathbb{E}_z \left[\cosh^{-4} \beta(\sqrt{2\bar{q}} + h)\right] > \frac{1}{2} \tag{5.19}$$

then the pressure of the SK model is strictly smaller than the replica symmetric pressure, namely:

$$\lim_{N \to \infty} p_N^{SK}(\beta, h) < p_{RS}(\beta, h; \bar{q}) \tag{5.20}$$

Proof. From Guerra's replica symmetry breaking bound we know that:

$$p_N^{SK}(\beta, h) \leq \inf \mathcal{P}(\beta, h; x(q)) \tag{5.21}$$

We only need to find a particular order parameter $\tilde{x}(q)$ such that $\mathcal{P}(\beta, h; \tilde{x}(q)) < p_{RS}(\beta, h; \bar{q})$ when (5.19) is fulfilled. The simplest choice we can make is:

$$x(q) = \begin{cases} 0 & \text{if } q \in [0, \bar{q}] \\ m & \text{if } q \in [\bar{q}, r] \\ 1 & \text{if } q \in [r, 1] \end{cases} \tag{5.22}$$

where obviously $m \in [0,1]$ and $r \in [\bar{q},1]$. With these sequences the Parisi functional becomes the 1-step RSB functional, described also in [10]:

$$\mathcal{P}(\beta,h;m,r) = \frac{\beta^2}{2}\left(1+m\bar{q}^2+(1-m)r^2-2r\right)+\log 2+ \frac{1}{m}\mathbb{E}_{z'}\left[\log \mathbb{E}_z \cosh^m \beta\left(\sqrt{2(r-\bar{q})}z+\sqrt{2\bar{q}}z'+h\right)\right] \quad (5.23)$$

When we take $m = 1$, and $r = \bar{q}$ and we go back to replica symmetric pressure, namely: $\mathcal{P}(\beta,h;1,\bar{q}) = p_{RS}(\beta,h)$, because we are at the zero-th step of the replica symmetry breaking. If $\mathcal{P}(\beta,h;m,r)$ arrives in $m = 1$ with positive derivative with respect to m for some r the proof is finished.

$$K(\beta,h;r) = \left.\frac{\partial\mathcal{P}(\beta,h;m,r)}{\partial m}\right|_{m=1} = -\frac{\beta^2}{2}(r^2-\bar{q}^2)- \\ -\mathbb{E}_{z'}\left[\log \mathbb{E}_z \cosh^m \beta\left(\sqrt{2(r-\bar{q})}z+\sqrt{2\bar{q}}z'+h\right)\right]+ \\ +\mathbb{E}_{z'}\left[\frac{\mathbb{E}_z \cosh\beta\left(\sqrt{2(r-\bar{q})}z+\sqrt{2\bar{q}}z'+h\right)\log\cosh\beta\left(\sqrt{2(r-\bar{q})}z+\sqrt{2\bar{q}}z'+h\right)}{\mathbb{E}_z \cosh\beta\left(\sqrt{2(r-\bar{q})}z+\sqrt{2\bar{q}}z'+h\right)}\right] \quad (5.24)$$

In order to establish its sign in a neighborhood of $\bar{q}$ we perform a Taylor expansion that yields:

$$K(\beta,h;\bar{q}) = 0 \quad (5.25)$$

$$\left.\frac{\partial K(\beta,h;r)}{\partial r}\right|_{r=\bar{q}} = 0 \quad (5.26)$$

$$\left.\frac{\partial^2 K(\beta,h;r)}{\partial r^2}\right|_{r=\bar{q}} = -\beta^2\left(1-2\beta^2\mathbb{E}_z\cosh^{-4}\beta\left(\sqrt{2\bar{q}}z+h\right)\right) \quad (5.27)$$

The latter is positive, at least for r in a neighborhood of $\bar{q}$, under the hypothesis (5.19). This concludes the proof. □

This leads us to define the AT line as follows:

$$\beta^2\mathbb{E}_z\cosh^{-4}\beta\left(\sqrt{2\bar{q}}z+h\right) = \frac{1}{2} \quad (5.28)$$

The AT line is conjectured to be the true line that separates the RS and full RSB regions in the phase space (β,h), though it has not been proved yet. We will adopt this assumption from now on.

Remark 5.1.2. It will be useful to notice that the point $(\beta, h) = (1/\sqrt{2}, 0)$ belongs to the line. In fact, for $\beta^2 \leq 1/2$, as discussed earlier, the consistency equation has only the trivial solution $\bar{q} = 0$, hence:

$$\beta^2 \mathbb{E}_z \cosh^{-4} \beta \left(\sqrt{2\bar{q}} z + h \right) = \frac{1}{2} \mathbb{E}_z \cosh^{-4}(0) = \frac{1}{2} \tag{5.29}$$

With the previous arguments we have just provided a proof for an important fact:

Proposition 5.1.6. *Assuming the AT line conjecture, for the SK model with vanishing external field, the annealed and replica symmetric regions coincide.*

5.1.2 A glance at the SK multi-species model

Similar statements hold for multi-species elliptic models, they are shown in [7]. For what follows, the temperature has been re-absorbed in the covariances. It is important to stress that the following theorem is only a sufficient condition, hence it does not define the annealed region properly, but only a subset of it.

Theorem 5.1.7. *Assume the same notations used for multi-species models. If the following conditions hold:*

$$\Delta^2 > 0 \tag{5.30}$$

$$\hat{\Delta}^2 = (\Delta^2)^{-1} - 2\alpha^{-1} > 0 \quad \textit{with } \alpha = diag(\alpha_1, \ldots, \alpha_K) \tag{5.31}$$

then the pressure of the model coincides with the annealed one:

$$p_N^{dms} = \frac{1}{N} \mathbb{E} \log \sum_\sigma e^{-H_N^{dms}(\sigma)} = \frac{1}{N} \log \sum_\sigma \mathbb{E} e^{-H_N^{dms}(\sigma)} = \log 2 + \frac{1}{2} \langle \tilde{\mathbf{1}}, \Delta^2 \tilde{\mathbf{1}} \rangle \tag{5.32}$$

Proof. The proof follows the same steps of that of Theorem 5.1.2. The upper bound is easily obtained by Jensen's inequality. Recalling the Hamiltonian gaussian family (4.2) and its covariance (4.3):

$$p_N^{dms} \leq \frac{1}{N} \log \mathbb{E} Z_N^{dms} = \frac{1}{N} \log e^{\frac{N}{2} \langle \tilde{\mathbf{q}}_{\sigma\sigma}, \Delta^2 \tilde{\mathbf{q}}_{\sigma\sigma} \rangle} = \log 2 + \frac{1}{2} \langle \tilde{\mathbf{1}}, \Delta^2 \tilde{\mathbf{1}} \rangle \tag{5.33}$$

The expectation over the disorder of ${Z_N^{dms}}^2$ yields:

$$\begin{aligned} \mathbb{E}[{Z_N^{dms}}^2] &= \sum_{\sigma,\tau} \mathbb{E} e^{-[H_\sigma^{dms} + H_\tau^{dms}]} = \sum_{\sigma,\tau} e^{\frac{N}{2}(2\langle \tilde{\mathbf{1}}, \Delta^2 \tilde{\mathbf{1}} \rangle + 2\langle \tilde{\mathbf{q}}_{\sigma\tau}, \Delta^2 \tilde{\mathbf{q}}_{\sigma\tau} \rangle)} = \\ &= e^{\frac{N}{2} 2\langle \tilde{\mathbf{1}}, \Delta^2 \tilde{\mathbf{1}} \rangle} 2^N \sum_\sigma e^{N \langle \tilde{\mathbf{m}}_\sigma, \Delta^2 \tilde{\mathbf{m}}_\sigma \rangle} = 2^{-N} \mathbb{E}^2[Z_N] \sum_\sigma e^{N \langle \tilde{\mathbf{m}}_\sigma, \Delta^2 \tilde{\mathbf{m}}_\sigma \rangle} \end{aligned} \tag{5.34}$$

where we have used again the gauge symmetry $\sigma_i \longrightarrow \sigma_i \tau_i$. Let us linearize the quadratic term at the exponent with Lemma (3.2.7).

$$
\begin{aligned}
\mathbb{E}[Z_N^{dms\,2}] = 2^{-N}\mathbb{E}^2[Z_N] \int_{\mathbb{R}^K} \frac{d^K x}{\sqrt{(2\pi)^K \det \Delta^2}} \sum_\sigma e^{-\frac{1}{2}\langle \mathbf{x}, (\Delta^2)^{-1}\mathbf{x}\rangle + \langle \tilde{\mathbf{m}}_\sigma, \sqrt{2N}\mathbf{x}\rangle} = \\
\mathbb{E}^2[Z_N] \int_{\mathbb{R}^K} \frac{d^K x}{\sqrt{(2\pi)^K \det \Delta^2}} \sum_\sigma e^{-\frac{1}{2}\langle \mathbf{x}, (\Delta^2)^{-1}\mathbf{x}\rangle + \sum_{s=1}^K N_p \log\cosh\left(\frac{\sqrt{2N}}{N_s}\mathbf{x}\right)} \leq \\
\leq \mathbb{E}^2[Z_N] \int_{\mathbb{R}^K} \frac{d^K x}{\sqrt{(2\pi)^K \det \Delta^2}} \sum_\sigma e^{-\frac{1}{2}\langle \mathbf{x}, \hat{\Delta}^2 \mathbf{x}\rangle} \qquad (5.35)
\end{aligned}
$$

The inequality follows from $\log\cosh x \leq x^2/2$ and $\hat{\Delta}^2$ is exactly the modified covariance matrix in the statement. Hence, if it is positive definite, the integral converges and the hypothesis of Lemma 5.1.1 are valid.

The rest of the proof is identical to that of Theorem 5.1.2. □

5.2 The Deep Boltzmann Machine (DBM)

In this section we present some new interesting ideas in [4]. By Deep Boltzmann Machine we mean a multi-layer SK model, which is a particular instance of hyperbolic multi-species model. This very characteristic makes it impossible to solve it exactly with standard methods, but something interesting can still be said on its annealed and replica symmetric phase space regions. The relation between the two, discussed later, is very important for Machine Learning purposes. In fact, these machines can be properly trained only out of the annealed region.

In order to keep it the smallest possible, we will consider the form factors α_l, defined in the previous chapters, as free parameters and then fix them appropriately. Hence, our phase space, in this case, will be richer.

The model is defined on a graph identical to that of a multi-layer deterministic CW model, except for the interactions that are extracted from Gaussian distributions, as described in what follows.

Definition 5.2.1 (DBM hamiltonian or cost function)**.** Let $\alpha_p = \frac{N_p}{N}$, with the same notations used for multi-species disordered and non disordered models. Further, let $J_{ij}^{(p)} \overset{\text{iid}}{\sim} \mathcal{N}(0, \Delta_p^2)$. The hamiltonian of the multi-layer SK model is:

$$
H_N^{DBM}(\sigma, J; \alpha) = -\frac{\sqrt{2}}{\sqrt{N}} \sum_{p=1}^{K-1} \sum_{(i,j)\in L_p \times L_{p+1}} J_{ij}^{(p)} \sigma_i \sigma_j \qquad (5.36)
$$

Remark 5.2.1. Observe that the model could have been equivalently defined through the covariances:

$$\mathbb{E}[H_\sigma^{DBM} H_\tau^{DBM}] = 2N \sum_{p=1}^{K-1} \Delta_p^2 \alpha_p \alpha_{p+1} q_p(\sigma,\tau) q_{p+1}(\sigma,\tau) = C^{DBM}(\sigma,\tau) \quad (5.37)$$

where the notations for the overlaps have been previously defined when we dealt with disordered multi-species models. The factor 2 or $\sqrt{2}$ in (5.36) is conventional and introduced for the sake of convenience.

5.2.1 Lower bound for the quenched pressure

The idea is again to use the interpolation method, interpolating the system with K decoupled SK models with an appropriate temperature. We will identify a rest with definite sign and this will give us the lower bound.

Inspired by the trick we used for the deterministic version, we define the following covariances:

$$\tilde{\Delta^2}_p(a) = \frac{\Delta_{p-1}^2}{a_{p-1}} + a_p \Delta_p^2 \quad \text{with } a_1, \dots, a_{K-1} > 0 \quad (5.38)$$

If the labels of the covariances and parameters are not in $[0, K]$ then the corresponding covariance is zero, *e.g.*: $\Delta_0^2 = 0$ and $\tilde{\Delta^2}_1 = a_1 \Delta_1^2$.

We will denote by $H_{N_p}^{SK}(\sigma; J)$ the hamiltonian of an SK model with N_p particles and the usual covariance: $\mathbb{E}[H_\sigma^{SK} H_\tau^{SK}] = N_p q_p^2(\sigma,\tau)$. With the previous notations, we have a lower bound for the quenched pressure:

Theorem 5.2.1. *The quenched pressure $p_N^{DBM}(\beta)$ of the DBM is bounded from below as follows:*

$$p_N^{DBM}(\beta) \geq \sum_{p=1}^{K} \alpha_p p_{N_p}^{SK}\left(\sqrt{\alpha_p \tilde{\Delta}_p^2}\beta\right) - \frac{\beta^2}{2} \sum_{p=1}^{K} \alpha_p^2 \tilde{\Delta}_p^2 + \beta^2 \sum_{p=1}^{K-1} \Delta_p^2 \alpha_p \alpha_{p+1} \quad (5.39)$$

Consequently:

$$\liminf_{N\to\infty} p_N^{DBM}(\beta) \geq \sup_{a\in(0,\infty)^{K-1}} \left[\sum_{p=1}^{K} \alpha_p p^{SK}\left(\sqrt{\alpha_p \tilde{\Delta}_p^2}\beta\right) - \frac{\beta^2}{2} \sum_{p=1}^{K} \alpha_p^2 \tilde{\Delta}_p^2\right] + \\ + \beta^2 \sum_{p=1}^{K-1} \Delta_p^2 \alpha_p \alpha_{p+1} \quad (5.40)$$

Proof. To begin with, we need the following interpolating hamiltonian and pressure:

$$H_N(t) = \sqrt{t} H_N^{DBM}(\sigma) + \sqrt{1-t} \sum_{p=1}^{K} H_{N_p}^{SK}(\sigma) \sqrt{\alpha_p \tilde{\Delta}_p^2} \tag{5.41}$$

$$p_N(t) = \frac{1}{N} \mathbb{E} \log \sum_{\sigma \in \Sigma_N} e^{-\beta H_N(t)} \tag{5.42}$$

$$p_N(0) = \frac{1}{N} \sum_{p=1}^{K} \mathbb{E} \log \sum_{\sigma \in \Sigma_{N_p}} e^{-\beta H_{N_p}^{SK}(\sigma)} = \sum_{p=1}^{K} \alpha_p p_{N_p}^{SK} \left(\beta \sqrt{\alpha_p \tilde{\Delta}_p^2} \right) \tag{5.43}$$

$$p_N(1) = p_N^{DBM}(\beta) \tag{5.44}$$

The first derivative of the interpolating pressure, after the integration by parts has been performed, is:

$$\begin{aligned} p_N'(t) = \frac{\beta^2}{2N} \mathbb{E} \Omega_{N,t}^{(2)} \left[C^{DBM}(\sigma, \sigma) - C^{DBM}(\sigma, \tau) - \sum_{p=1}^{K} \alpha_p \tilde{\Delta}_p^2 (N_p - N_p q_p^2(\sigma, \tau)) \right] = \\ = \beta^2 \sum_{p=1}^{K-1} \Delta_p^2 \alpha_p \alpha_{p+1} - \frac{\beta^2}{2} \sum_{p=1}^{K} \alpha_p^2 \tilde{\Delta}_p^2 + \frac{\beta^2}{2} R_N(\alpha, a, \mathbf{q}) \end{aligned} \tag{5.45}$$

We already recognize the first terms that appear in the lower bound. Now the goal is to proof that the last term has a definite sign:

$$\begin{aligned} R_N(\alpha, a, \mathbf{q}) = -2 \sum_{p=1}^{K-1} \Delta_p^2 \alpha_p \alpha_{p+1} q_p(\sigma, \tau) q_{p+1}(\sigma, \tau) + \sum_{p=1}^{K} \alpha_p^2 \left(\frac{\Delta_{p-1}^2}{a_{p-1}} + a_p \Delta_p^2 \right) q_p^2(\sigma, \tau) = \\ = \sum_{p=1}^{K-1} \Delta_p^2 \left(\alpha_p \sqrt{a_p} q_p(\sigma, \tau) - \frac{\alpha_{p+1}}{\sqrt{a_p}} q_{p+1}(\sigma, \tau) \right)^2 \geq 0 \end{aligned} \tag{5.46}$$

Finally, the statement follows from a simple application of the theorem of integral calculus. □

5.2.2 Annealing in the DBM

The form of the annealed pressure is a bit different for such a model, due to the peculiar form of the convariances (5.37).

Definition 5.2.2 (Annealed pressure of DBM)**.** The annealed pressure of a Deep Boltzmann Machine, whose hamiltonians have covariances (5.37) is:

$$p^A(\beta,\alpha) = \frac{1}{N}\log\sum_{\sigma}\mathbb{E}e^{-\beta H_\sigma^{DBM}} = \log 2 + \beta^2\sum_{p=1}^{K-1}\Delta_p^2\alpha_p\alpha_{p+1} \tag{5.47}$$

Remark 5.2.2. As usual, the concavity of log and Jensen's inequality imply that the quenched pressure is bounded from above by p^A.

If we have a look at the lower bound in (5.39), it is reasonable to expect that the annealed pressure and the quenched one will coincide when each of the interpolated SK models is in its own annealed region. This leads us to define the following subset of the phase space:

$$A_K = \left\{(\beta,\alpha):\ \alpha_p\tilde{\Delta}_p^2(a)\beta^2 \le \frac{1}{2}\text{ for some } a\in(0,\infty)^{K-1}\right\} \tag{5.48}$$

Theorem 5.2.2 (Annealing condition, DBM)**.** *If $(\beta,\alpha)\in A_K$ then the quenched pressure and the annealed one coincide in the thermodynamic limit:*

$$\lim_{N\to\infty} p_N^{DBM}(\beta,\alpha) = p^A(\beta,\alpha) \tag{5.49}$$

Proof. The lower bound in (5.39) can be rewritten in terms of the annealed pressure of the SK models:

$$\liminf_{N\to\infty} p_N^{DBM}(\beta,\alpha) \ge \sup_{a\in(0,\infty)^{K-1}}\sum_{p=1}^{K}\alpha_p\left[p^{SK}\left(\sqrt{\alpha_p\tilde{\Delta}_p^2}\beta\right) - p^{an}\left(\sqrt{\alpha_p\tilde{\Delta}_p^2}\beta\right)\right] + \\ + \log 2 + \beta^2\sum_{p=1}^{K-1}\Delta_p^2\alpha_p\alpha_{p+1} = p^A(\beta,\alpha) \tag{5.50}$$

As predicted, the terms in the square brackets cancel each other, leaving a lower bound containing only the annealed pressure of the DBM. The annealed pressure is always an upper bound for the quenched pressure, and this concludes the proof. □

The region (5.48) may be rewritten as done in the following proposition.

Proposition 5.2.3. *For $K = 2, 3, 4$ we have:*

$$A_K = \left\{(\beta, \alpha) : 4\beta^4 \leq \phi_K(\alpha)\right\} \tag{5.51}$$

$$\phi_2(\alpha) = \frac{1}{\Delta_1^4 \alpha_1 \alpha_2} \tag{5.52}$$

$$\phi_3(\alpha) = \frac{1}{\Delta_1^4 \alpha_1 \alpha_2 + \Delta_2^4 \alpha_2 \alpha_3} \tag{5.53}$$

$$\phi_4(\alpha) = \min\left\{t > 0 : 1 - t(\Delta_1^4 \alpha_1 \alpha_2 + \Delta_2^4 \alpha_2 \alpha_3 + \Delta_3^4 \alpha_3 \alpha_4) + t^2 \Delta_1^4 \Delta_3^4 \alpha_1 \alpha_2 \alpha_3 \alpha_4 = 0\right\} \tag{5.54}$$

Proof. For the sake of clarity, the proof has been divided in the three cases.

Two layers: $K = 2$ By the definition in (5.48), $(\beta, \alpha_1, \alpha_2) \in A_2$ iff for some $a_1 \in (0, \infty)$:

$$\begin{cases} 2\alpha_1 \Delta_1^2 \beta^2 a_1 \leq 1 \\ \frac{2\alpha_2 \Delta_1^2 \beta^2}{a_1} \leq 1 \end{cases} \quad \Leftrightarrow \quad 4\beta^4 \Delta_1^4 \alpha_1 \alpha_2 \leq 1 \tag{5.55}$$

Three layers: $K = 3$ Again $(\beta, \alpha_1, \alpha_2, \alpha_3) \in A_3$ iff for some $a_1, a_2 \in (0, \infty)$:

$$\begin{cases} 2\alpha_1 \Delta_1^2 \beta^2 a_1 \leq 1 \\ 2\beta^2 \alpha_2 \tilde{\Delta}_2^2(a_1, a_2) \leq 1 \\ \frac{2\alpha_3 \Delta_2^2 \beta^2}{a_2} \leq 1 \end{cases} \tag{5.56}$$

$\tilde{\Delta}_2^2(a_1, a_2)$ is decreasing in a_1 and increasing in a_2. Hence, *w.l.o.g.*:

$$\tilde{\Delta}_2^2\left(\frac{1}{2\alpha_1 \Delta_1^2 \beta^2}, 2\alpha_3 \Delta_2^2 \beta^2\right) \leq \frac{1}{2\alpha_2 \beta^2} \quad \Leftrightarrow \quad 4\beta^4(\alpha_1 \alpha_2 \Delta_1^4 + \alpha_2 \alpha_3 \Delta_2^4) \leq 1 \tag{5.57}$$

Four layers: $K = 4$ $(\beta, \alpha_1, \alpha_2, \alpha_3, \alpha_4) \in A_4$ iff for some $a_1, a_2, a_3 \in (0, \infty)$:

$$\begin{cases} 2\alpha_1 \Delta_1^2 \beta^2 a_1 \leq 1 \\ 2\beta^2 \alpha_2 \tilde{\Delta}_2^2(a_1, a_2) \leq 1 \\ 2\beta^2 \alpha_3 \tilde{\Delta}_3^2(a_2, a_3) \leq 1 \\ \frac{2\alpha_4 \Delta_2^2 \beta^2}{a_3} \leq 1 \end{cases} \tag{5.58}$$

Again, using the motonicity properties above, we can rewrite the annealing condition in this way:

$$\exists a_2 > 0 \ : \quad \begin{cases} 2\beta^2\alpha_2\tilde{\Delta}_2^2\left(\frac{1}{2\beta^2\alpha_1\Delta_1^2}, a_2\right) \leq 1 \\ 2\beta^2\alpha_3\tilde{\Delta}_3^2\left(a_2, 2\beta^2\alpha_4\Delta_3^2\right) \leq 1 \end{cases} \tag{5.59}$$

or equivalently:

$$\begin{cases} (1 - 4\beta^4\Delta_1^4\alpha_1\alpha_2)(1 - 4\beta^4\Delta_3^4\alpha_3\alpha_4) \geq 4\beta^4\Delta_2^4\alpha_2\alpha_3 \\ 1 - 4\beta^4\Delta_1^4\alpha_1\alpha_2 \geq 0 \end{cases} \tag{5.60}$$

If we set $t = 4\beta^4$ in the first inequality, we would get $t \leq t_- \vee t \geq t_+$, with $t_\pm$ such that $1 - t_\pm(\Delta_1^4\alpha_1\alpha_2 + \Delta_2^4\alpha_2\alpha_3 + \Delta_3^4\alpha_3\alpha_4) + t_\pm^2\Delta_1^4\Delta_3^4\alpha_1\alpha_2\alpha_3\alpha_4 = 0$. It is not difficult to prove that $t_- \leq 1/(\Delta_1^4\alpha_1\alpha_2) \leq t_+$, and this implies that the second inequality in the system selects only $t \leq t_-$. This concludes the proof. □

As anticipated in the introduction, we want to minimize, in some sense, the annealing region. In order to do this, we set the parameters α in the tightest possible way, *i.e.* minimizing $\phi_K(\alpha)$ with the constraint $\sum_{p=1}^4 \alpha_p = 1$. For $K = 2$ we obtain:

$$\alpha_1 = \alpha_2 = \frac{1}{2} \tag{5.61}$$

Things get rapidly complicated as the number of layers increases. An interesting thing arises in the computation for $K = 3$: the two variances must be equal in order to have a solution to the Lagrange minimization method.

$$\frac{\partial\Lambda}{\partial\alpha_1} = -\phi_3^2(\alpha)\Delta_1^4\alpha_2 - \lambda = 0 \tag{5.62}$$

$$\frac{\partial\Lambda}{\partial\alpha_2} = -\phi_3^2(\alpha)(\Delta_1^4\alpha_1 + \Delta_2^4\alpha_3) - \lambda = 0 \tag{5.63}$$

$$\frac{\partial\Lambda}{\partial\alpha_3} = -\phi_3^2(\alpha)\Delta_2^4\alpha_2 - \lambda = 0 \tag{5.64}$$

where $\Lambda = \phi_3(\alpha) - \lambda(\alpha_1 + \alpha_2 + \alpha_3 - 1)$. The first one, together with the third one, imply that either $\alpha_2 = 0$ or $\Delta_1^2 = \Delta_2^2$. $\alpha_2 = 0$ is not possible, because it would violate the second equation. Hence we are forced to admit $\Delta_1^2 = \Delta_2^2$. In that case:

$$\alpha_2 = \frac{1}{2}, \quad \alpha_1 + \alpha_3 = \frac{1}{2} \tag{5.65}$$

The fact that the Lagrange method fails tells us that the extremal point is on the frontier of the simplex of the possible α's. Therefore, we have to perform a minimization with $\alpha_1 = 0$ or $\alpha_3 = 0$. $\alpha_2 = 0$ again is not possible, because it would bring ϕ_3 up to infinity. For example, if $\alpha_3 = 0$, then:

$$\alpha_1 = \alpha_2 = \frac{1}{2}, \quad \alpha_3 = 0 \tag{5.66}$$

Notice that this still satisfies (5.65). Here a clarification is needed: when a ratio α vanishes, it does not mean that we have zero neurons (spins) in the corresponding layer. It only means that the number of them grows sub-linearly with the number of particles in the system. Thus, in the thermodynamic limit their fraction becomes negligible, *i.e.* 0.

Finally, assuming equal variances and for $K = 4$, standard computations yield:

$$\alpha_1 = 0,\, \alpha_3 = \frac{1}{2},\, \alpha_2 + \alpha_4 = \frac{1}{2} \quad \text{or} \quad \alpha_4 = 0,\, \alpha_2 = \frac{1}{2},\, \alpha_3 + \alpha_1 = \frac{1}{2} \tag{5.67}$$

5.2.3 A possible replica symmetric pressure for DBM

Unfortunately, since the DBM is an hyperbolic model, it is not always possible to identify a remainder with definite sign with the interpolation technique. For this reason what follows relies only on the fact that in the replica symmetric region the overlaps do not fluctuate. Let us reintroduce, for the moment, the deterministic weights $W_N(\beta, \mathbf{h}; \sigma) = \exp\left(\beta \sum_{p=1}^{K} h_p \sum_{i\in L_p} \sigma_i\right)$.

Proposition 5.2.4. *The following sum rule holds:*

$$p_N^{DBM}(\beta, \mathbf{h}; \alpha) = p_{RS}^{DBM}(\beta, \mathbf{h}; \alpha, \mathbf{q}) - \int_0^1 dt\, \mathbb{E}\Omega_{N,t}^{(2)}\left[R_N\right] \tag{5.68}$$

Where the replica symmetric pressure has been introduced:

$$\begin{aligned} p_{RS}^{DBM}(\beta, \mathbf{h}; \alpha, \mathbf{q}) &= \log 2 + \\ &+ \sum_{p=1}^{K} \alpha_p \mathbb{E}_z \log\cosh\beta\left(\sqrt{2}\sqrt{\alpha_{p-1}\Delta_{p-1}^2 q_{p-1} + \alpha_{p+1}\Delta_p^2 q_{p+1}} + h_p\right) + \\ &+ \beta^2 \sum_{p=1}^{K-1} \Delta_p^2 \alpha_p \alpha_{p+1}(1-q_p)(1-q_{p+1}) \end{aligned} \tag{5.69}$$

together with the remainder:

$$R_N(\sigma,\tau) = \sum_{p=1}^{K-1} \alpha_p \alpha_{p+1} (q_p(\sigma,\tau) - q_p)(q_{p+1}(\sigma,\tau) - q_{p+1}) \tag{5.70}$$

Remark 5.2.3. The choice of the interpolating hamiltonian is not casual. In fact, if we remember the choices made in (4.27), we see that the argument of the square root is exactly the projection on the layer p of the vector $\Delta^2\mathbf{q}$, where this time Δ^2 is tridiagonal with vanishing diagonal elements.

Proof. The strategy is again to interpolate the hamiltonian of the DBM with that of K decoupled free systems, with an appropriately modified temperature.

$$H_N(t) = \sqrt{t} H_N^{DBM}(\sigma) - \sqrt{1-t} \underbrace{\sum_{p=1}^{K} \sqrt{2} \sqrt{\alpha_{p-1}\Delta^2_{p-1} q_{p-1} + \alpha_{p+1}\Delta^2_p q_{p+1}} \sum_{i \in L_p} J_i^{(p)} \sigma_i}_{\tilde{H}_N(\sigma)} \tag{5.71}$$

where obviously: $\alpha_0 = \alpha_{K+1} = 0$ for convenience, and $J_i^{(p)} \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$. The family of non interacting models, with hamiltonian $\tilde{H}_N$ has the following covariance:

$$\begin{aligned}\tilde{C}(\sigma,\tau) = \mathbb{E}[\tilde{H}_\sigma \tilde{H}_\tau] = 2N \sum_{p=1}^{K} \alpha_p (\alpha_{p-1}\Delta^2_{p-1} q_{p-1} + \alpha_{p+1}\Delta^2_p q_{p+1}) q_p(\sigma,\tau) = \\ = 2N \sum_{p=1}^{K-1} \Delta^2_p \alpha_p \alpha_{p+1} (q_p q_{p+1}(\sigma,\tau) + q_{p+1} q_p(\sigma,\tau))\end{aligned} \tag{5.72}$$

The next step is the computation of the first derivative of the interpolating pressure. Through the integration by parts Lemma 2.1.1 we can write it in this form:

$$p'_N(t) = \frac{\beta^2}{2N} \mathbb{E}\Omega^{(2)}_{N,t} \left[C^{DBM}(\sigma,\sigma) - C^{DBM}(\sigma,\tau) - \tilde{C}(\sigma,\sigma) + \tilde{C}(\sigma,\tau) \right] \tag{5.73}$$

After a shift in the sums variable p one can rewrite the previous in this way:

$$\begin{aligned}p'_N(t) = \beta^2 \mathbb{E}\Omega^{(2)}_{N,t} \Bigg[\sum_{p=1}^{K-1} \alpha_p \alpha_{p+1} \Delta^2_p (1-q_p)(1-q_{p+1}) - \\ - \sum_{p=1}^{K-1} \alpha_p \alpha_{p+1} \Delta^2_p (q_p(\sigma,\tau) - q_p)(q_{p+1}(\sigma,\tau) - q_{p+1}) \Bigg]\end{aligned} \tag{5.74}$$

We clearly recognize the rest appearing in the statement in the last term. The interpolating pressure at the extremes is:

$$p_N(0) = \frac{1}{N}\mathbb{E}\log\sum_{\sigma\in\Sigma_N} W_N(\beta,\mathbf{h};\sigma)e^{-\beta\tilde{H}_N(\sigma)} =$$

$$= \log 2 + \sum_{p=1}^{K}\alpha_p\mathbb{E}_z\log\cosh\beta\left(\sqrt{2}\sqrt{\alpha_{p-1}\Delta_{p-1}^2 q_{p-1} + \alpha_{p+1}\Delta_p^2 q_{p+1}} + h_p\right) \quad (5.75)$$

whereas $p_N(1) = p_N^{DBM}$. Finally, a simple application of the theorem of integral calculus yields the result. □

The replica symmetric pressure for the DBM produces the following consistency equation:

$$q_p = \mathbb{E}_z\tanh^2\beta\left(\sqrt{2}\sqrt{\alpha_{p-1}\Delta_{p-1}^2 q_{p-1} + \alpha_{p+1}\Delta_p^2 q_{p+1}} + h_p\right) = F_p(\mathbf{q},\mathbf{h}) \quad (5.76)$$

Notice that for $h_p = 0\ \forall p = 1,\dots,K$, the point $\bar{q}_p = 0\ \forall p = 1,\dots,K$ is always a solution. In analogy to what discussed for the SK model let us find a sufficient condition for $\bar{q}_p = 0$ to be a stable stationary point.

It is sufficient to impose that the map on the r.h.s. is a contraction (for $h = 0$). This can be done by checking the spectral radius of its jacobian matrix:

$$\left.\frac{\partial F_p}{\partial q_{p'}}\right|_{\mathbf{q}=0} = 2\beta^2(\delta_{p-1,p'}\Delta_{p-1}^2\alpha_{p-1} + \delta_{p+1,p'}\Delta_p^2\alpha_{p+1}) = (JF(0,0))_{pp'} \quad (5.77)$$

in 0. The eigenvalues equations for $K = 2,3,4$ are listed below:

$$D_2 = \lambda^2 - 4\beta^4\Delta_1^4\alpha_1\alpha_2 = 0 \quad (5.78)$$

$$D_3 = \lambda^3 - 4\lambda\beta^4(\Delta_1^4\alpha_1\alpha_2 + \Delta_2^4\alpha_2\alpha_3) = 0 \quad (5.79)$$

$$D_4 = \lambda^4 - 4\lambda^2\beta^4(\Delta_1^4\alpha_1\alpha_2 + \Delta_2^4\alpha_2\alpha_3 + \Delta_3^4\alpha_3\alpha_4) + 16\beta^8\Delta_1^4\Delta_3^4\alpha_1\alpha_2\alpha_3\alpha_4 = 0 \quad (5.80)$$

with $D_K = det(\lambda\mathbb{1} - JF(0,0))$.

The following proposition collects the results, obtained with standard computations, for $K = 2,3,4$.

Proposition 5.2.5. *The jacobian matrix* $\left.\frac{\partial F_p}{\partial q_{p'}}\right|_{\mathbf{q}=0}$ *has a spectral radius smaller than one if and only if:* $4\beta^4 < \phi_K(\alpha)$.

Remark 5.2.4. By chance, the previous stability condition for the solution $q_p = 0$ is identical, except for the frontier, to the annealed condition for the DBM. In analogy with the SK model, this fact lets us think that A_K may be the entire annealing region for the DBM.

Conclusion and outlooks

To conclude, we briefly list the main achievements together with possible new interesting points we would like to pursue in a follow-up of this thesis.

In the first two chapters, the main tools to deal with mean field models are provided. In particular, the Guerra-Toninelli interpolation scheme turned out to be a useful technique for most of our proofs, since it relies only on the convexity of some functions involved in our computations. It is not explicitly shown in the present work, but interpolation can provide the replica symmetry breaking upper bound to the SK model pressure. We preferred to use the Aizenmann-Sims-Starr scheme, a more general approach, that allowed us to give a representation of the Parisi *ansatz* in terms of RPCs and cavity functional.

A real step towards the study of Boltzmann machines was taken in chapter 3 and chapter 4, where multi-species models were studied. In the deterministic case, the multi-layer Curie-Weiss model was solved by interpolating it with one body Hamiltonians, related to non-interacting systems. For disordered multi-species models, an adapted cavity functional was defined that led us to an extended variational principle for these models. The latter is a reorganization of the ideas presented in [7][19].

The content of the final chapter can be regarded as the main subject of this thesis, both for the model analysed in it, the DBM, and for the methods used, justifying in some sense the path followed up to it. However, chapter 5 is more the starting point of a new project than a true conclusion. There are many questions to be answered. Due to the lack of convexity of the interactions, we are not even able to prove that the thermodynamic limit of the DBM exists. Interpolation cannot help us in this case. Furthermore, the free energy in the thermodynamic limit is still unknown, we only have a candidate as an upper bound.

Nevertheless, some new perspectives arose from the study of the replica symmetric and annealed regions of the DBM. In particular, the knowledge of the latter turns out to be very important in machine learning applications, as it is mandatory to escape it. The intuitive reason is that in this very region the order parameter has only a vanishing trivial solution. Moreover, the replica method, within the replica symmetric *ansatz*, has been widely and successfully used to study the asymptotic behaviour of quenched free energies corresponding to the cost functions, or Hamiltonians in our jargon, involved in such learning algorithms.

The squeezing of the aforementioned annealed region has been performed only with four or less layers, but there is hope that with similar arguments one can deduce the optimal (from this point of view) form factors with a generic number of layers. It should be also stressed that the conditions we have found on the α's are just the ones that ensure a "little" annealed region, but no other properties of these deep network, such as the generalisation, are discussed here. Therefore, it would be interesting search for other possible geometric constraints on the architecture coming from thermodynamics, this time on the depth of the network, instead of the width of the layers.

Ringraziamenti

Il primo ringraziamento non può che andare al mio relatore, il prof. Pierluigi Contucci, per la fiducia che ha riposto in me in questi mesi assegnandomi dei temi importanti ed a me nuovi, ai quali mi sono veramente appassionato. Inutile dire che la sua supervisione è stata fondamentale per la realizzazione di questo lavoro nella sua interezza.

Voglio ringraziare anche Francesco Guerra che ha generosamente condiviso con noi nuovi spunti, dai quali poi questa tesi è nata.

Un enorme grazie va al mio correlatore, il dott. Emanuele Mingione, e al dott. Diego Alberici per la pazienza che hanno avuto, per i consigli ed anche per il piacevole tempo passato insieme.

Grazie di cuore anche alla mia famiglia. A mio padre Adriano, che riesce sempre a farmi sembrare possibile ogni impresa. A mia madre Sabrina, che ha sempre le parole giuste per tirarmi su di morale.

Un ringraziamento speciale va ad Arianna, che ha l'innata capacità di distrarmi quando serve e che mi tiene sempre coi piedi per terra.

Un affettuoso saluto, infine, ai miei amici del Collegio Superiore, coi quali ho passato cinque magnifici anni, e ad Andrea Calcinari, formidabile compagno di studi.

Bibliography

[1] M. Aizenman, J. L. Lebowitz, and D. Ruelle. "Some rigorous results on the Sherrington-Kirkpatrick spin glass model". In: *Communications in Mathematical Physics* 112.1 (Mar. 1987), pp. 3–20. DOI: 10.1007/BF01217677.

[2] Michael Aizenman, Robert Sims, and Shannon L. Starr. "Extended variational principle for the Sherrington-Kirkpatrick spin-glass model". In: 68.21, 214403 (Dec. 2003), p. 214403. DOI: 10.1103/PhysRevB.68.214403. arXiv: cond-mat/0306386 [cond-mat.dis-nn].

[3] Michael Aizenman, Robert Sims, and Shannon L. Starr. "Mean-Field Spin Glass models from the Cavity–ROSt Perspective". In: *arXiv e-prints*, math-ph/0607060 (July 2006), math–ph/0607060. arXiv: math-ph/0607060.

[4] Diego Alberici et al. "Annealing and replica-symmetry in Deep Boltzmann machines". In: *in preparation* (Sept. 2019).

[5] Diego Alberici et al. "Finite-size corrections for the attractive mean-field monomer-dimer model". In: *Journal of Physics A: Mathematical and Theoretical* 52.10 (Feb. 2019), p. 105001. DOI: 10.1088/1751-8121/ab01ab.

[6] Adriano Barra et al. "Mean field bipartite spin models treated with mechanical techniques". In: *European Physical Journal B* 87, 74 (Mar. 2014), p. 74. DOI: 10.1140/epjb/e2014-40952-4. arXiv: 1310.5901 [cond-mat.dis-nn].

[7] Adriano Barra et al. "Multi-Species Mean Field Spin Glasses. Rigorous Results". In: *Annales Henri Poincaré* 16.3 (Mar. 2015), pp. 691–708. ISSN: 1424-0661. DOI: 10.1007/s00023-014-0341-5. URL: https://doi.org/10.1007/s00023-014-0341-5.

[8] A. Bianchi, P. Contucci, and C. Giardina'. "Thermodynamic Limit for Mean-Field Spin Models". In: *arXiv e-prints*, math-ph/0311017 (Nov. 2003), math–ph/0311017. arXiv: math-ph/0311017 [math-ph].

[9] Anton Bovier. *Statistical Mechanics of Disordered Systems: A Mathematical Perspective*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2006. DOI: 10.1017/CBO9780511616808.

[10] Pierluigi Contucci and Cristian Giardinà. *Perspectives on Spin Glasses*. Cambridge University Press, 2012. DOI: 10.1017/CBO9781139049306.

[11] Pierluigi Contucci, Sandro Graffi, and Stefano Isola. "Mean Field Behaviour of Spin Systems with Orthogonal Interaction Matrix". In: *Journal of Statistical Physics* 106.5 (Mar. 2002), pp. 895–914. ISSN: 1572-9613. DOI: 10.1023/A:1014073417247.

[12] Micaela Fedele and Pierluigi Contucci. "Scaling Limits for Multi-species Statistical Mechanics Mean-Field Models". In: *Journal of Statistical Physics* 144.6 (Sept. 2011), pp. 1186–1205. DOI: 10.1007/s10955-011-0334-4. arXiv: 1011.3216.

[13] Ignacio Gallo and Pierluigi Contucci. "Bipartite Mean Field Spin Systems. Existence and Solution". In: *arXiv e-prints* (Oct. 2007). arXiv: 0710.0800 [cond-mat.stat-mech].

[14] Francesco Guerra. "Broken Replica Symmetry Bounds in the Mean Field Spin Glass Model". In: *Communications in Mathematical Physics* 233.1 (Feb. 2003), pp. 1–12. ISSN: 1432-0916. DOI: 10.1007/s00220-002-0773-5. URL: https://doi.org/10.1007/s00220-002-0773-5.

[15] Francesco Guerra. "Fluctuations and thermodynamic variables in mean field spin glass models". In: *arXiv e-prints* (Dec. 2012), arXiv:1212.2905. arXiv: 1212.2905 [cond-mat.dis-nn].

[16] Francesco Guerra. *Spin Glasses*. In: (Feb. 2008). URL: arXiv:cond-mat/0507581.

[17] Francesco Guerra and Fabio Lucio Toninelli. "The Thermodynamic Limit in Mean Field Spin Glass Models". In: *Communications in Mathematical Physics* 230.1 (Sept. 2002), pp. 71–79. DOI: 10.1007/s00220-002-0699-y.

[18] D. Panchenko. *The Sherrington-Kirkpatrick Model*. Springer Monographs in Mathematics. Springer New York, 2015. URL: https://books.google.it/books?id=qJJOrgEACAAJ.

[19] Dmitry Panchenko. "The free energy in a multi-species Sherrington–Kirkpatrick model". In: *Ann. Probab.* 43.6 (Nov. 2015), pp. 3494–3513. DOI: 10.1214/14-AOP967. URL: https://doi.org/10.1214/14-AOP967.

[20] G Parisi. "A sequence of approximated solutions to the S-K model for spin glasses". In: *Journal of Physics A: Mathematical and General* 13.4 (Apr. 1980), pp. L115–L121. DOI: 10.1088/0305-4470/13/4/009. URL: https://doi.org/10.1088%2F0305-4470%2F13%2F4%2F009.

[21] G. Parisi. "Infinite Number of Order Parameters for Spin-Glasses". In: *Phys. Rev. Lett.* 43 (23 Dec. 1979), pp. 1754–1756. DOI: 10.1103/PhysRevLett.43.1754. URL: https://link.aps.org/doi/10.1103/PhysRevLett.43.1754.

[22] Giorgio Parisi. "Order Parameter for Spin-Glasses". In: *Phys. Rev. Lett.* 50 (24 June 1983), pp. 1946–1948. DOI: 10.1103/PhysRevLett.50.1946. URL: https://link.aps.org/doi/10.1103/PhysRevLett.50.1946.

[23] David Sherrington and Scott Kirkpatrick. "Solvable Model of a Spin-Glass". In: *Phys. Rev. Lett.* 35 (26 Dec. 1975), pp. 1792–1796. DOI: 10.1103/PhysRevLett.35.1792. URL: https://link.aps.org/doi/10.1103/PhysRevLett.35.1792.

[24] Michel Talagrand. *Mean Field Models for Spin Glasses. Volume I: Basic Examples*. Vol. 54. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics. Springer-Verlag Berlin Heidelberg, 2011. DOI: 10.1007/978-3-642-15202-3.

[25] Michel Talagrand. *Spin Glasses: A Challenge for Mathematicians. Cavity and Mean Field Models*. Vol. 46. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics. Springer-Verlag Berlin Heidelberg, 2003.

[26] Michel Talagrand. "The Parisi Formula". In: *Annals of Mathematics* 163.1 (2006), pp. 221–263. ISSN: 0003486X. URL: http://www.jstor.org/stable/20159953.

[27] F.L. Toninelli. "About the Almeida-Thouless transition line in the Sherrington-Kirkpatrick mean-field spin glass model". In: *EPL (Europhysics Letters)* 60.5 (Dec. 2002), pp. 764–767. DOI: 10.1209/epl/i2002-00667-5. arXiv: cond-mat/0207296.